ZAKARIA COULIBALY

+1 (215) 535 9838 | zcoulibalyeng@gmail.com | github.com/levisstrauss | linkedin.com/in/codemon

SUMMARY

Master's student in Computer Science at UIUC with hands-on experience building end-to-end ML pipelines, from experimentation to production deployment. Skilled in AWS ML infrastructure (SageMaker, Lambda, Step Functions), MLOps practices, and statistical model validation. Seeking an applied ML or data science role to drive business impact through rigorous experimentation and scalable ML systems.

TECHNICAL SKILLS

Languages: Python, SQL, Java, C/C++, JavaScript, TypeScript

Data & ML: PyTorch, Scikit-learn, Pandas, NumPy, MLflow, Weights & Biases

Cloud Infrastructure: AWS (SageMaker, Lambda, Step Functions, Bedrock),Terraform, Docker, Redis, kubernetes PostgreSQL

Backend & APIs: FastAPI, REST APIs, CI/CD Pipelines, GitHub Actions

TECHNICAL PROJECTS

ML Workflow for Image Classification – AWS MLOps Pipeline
AWS SageMaker, Lambda, Step Functions, Python
- Designed end-to-end ML pipeline on AWS with automated training, deployment, and inference using Step Functions
- Deployed model endpoints with Lambda functions for serverless, scalable image classification API
- Implemented CI/CD workflow for model retraining and monitoring in production environment

NYC Airbnb Price Prediction – MLOps Pipeline
Python, MLflow, Weights & Biases, Hydra, Scikit-learn
- Built end-to-end ML pipeline with automated data versioning, experiment tracking, and model registry using MLflow and W&B
- Implemented data validation with statistical drift detection (KL divergence) for reliable weekly retraining cycles
- Configured hyperparameter optimization sweeps using Hydra, enabling reproducible experiments and versioned releases

Enterprise RAG System for Technical Documentation Search System
AWS Bedrock, Terraform, Aurora PostgreSQL, Python, Streamlit
- Architected scalable search application using AWS serverless infrastructure, enabling instant retrieval of technical specifications across 5 product lines
- Implemented infrastructure-as-code with Terraform managing VPC isolation, database configuration, S3 storage, and automated CI/CD deployment pipeline
- Designed vector similarity search (HNSW indexing) with query validation layer, achieving 100% accuracy on specification lookups

Census Income Classification – Full-Stack ML Deployment
Python, FastAPI, Scikit-learn, Streamlit, GitHub Actions, Render
- Deployed production ML pipeline with FastAPI REST API, CI/CD automation via GitHub Actions, and cloud hosting on Render
- Implemented slice-based model evaluation across 8 demographic features to detect performance disparities and potential bias
- Built interactive Streamlit frontend with real-time predictions, achieving 79.7% precision on income classification task

EDUCATION

Master of Computer Science
University of Illinois Urbana-Champaign | **GPA: 3.94** | Expected: **Fall 2026**
- Relevant Coursework: Deep Learning for Healthcare, Natural Language Processing, Applied Machine Learning, Distributed Systems, Database Systems, Theory and Practice of Data Cleaning

Bachelor of Science in Computer Science & Minor in Mathematics
Penn State University | May 2024

PROFESSIONAL DEVELOPMENT

- Building GenAI with Amazon Bedrock - Udacity
- Agentic AI Nanodegree - Udacity
- AI Programming with Python Nanodegree - Udacity
- Machine Learning DevOps Engineer - Udacity
- AWS Machine Learning Fundamentals Nanodegree - Udacity
- Machine Learning Specialization - Stanford University & DeepLearning.AI