Marcus Williams

marcus.williams@email.com | (512) 555-0187 | Austin, TX | linkedin.com/in/marcuswilliams | github.com/mwilliams

PROFESSIONAL SUMMARY

AI Engineer with 3 years of experience building and deploying intelligent automation solutions. Proven track record of developing production-ready AI agents that improved operational efficiency by 40%. Expertise in LLM integration, prompt engineering, and building scalable ML pipelines. Strong background in transforming business requirements into technical solutions.

EXPERIENCE

AI Engineer

CloudTech Solutions | January 2022 - Present

- Developed and deployed 3 production AI agents for document processing, reducing manual review time by 40%
- Implemented RAG (Retrieval-Augmented Generation) system using Pinecone and OpenAI embeddings for internal knowledge base queries
- Built automated email response system using GPT-4 that handles 1,000+ customer inquiries daily with 92% satisfaction rate
- Created monitoring dashboard for AI agent performance tracking and error detection
- Optimized prompt templates and fine-tuned models, improving response accuracy from 78% to 91%
- Collaborated with product team to define AI capabilities and user experience flows

Machine Learning Engineer

DataDriven Inc | June 2020 - December 2021

- Built classification models for fraud detection achieving 94% precision
- Deployed ML models using FastAPI and Docker on AWS infrastructure
- Implemented CI/CD pipelines for model training and deployment automation
- Conducted A/B testing to evaluate model performance in production

EDUCATION

Master of Science in Artificial Intelligence - Georgia Institute of Technology
GPA: 3.8/4.0 | Graduated: December 2021

Bachelor of Science in Computer Engineering - University of Texas at Austin
GPA: 3.6/4.0 | Graduated: May 2019

TECHNICAL SKILLS

Programming Languages: Python, JavaScript, Go, SQL
AI/ML Frameworks: LangChain, LlamaIndex, OpenAI API, Anthropic API, HuggingFace, PyTorch, TensorFlow
Tools & Platforms: n8n, Make.com, Docker, Kubernetes, AWS (SageMaker, Lambda, EC2), PostgreSQL
Vector Databases: Pinecone, Weaviate, ChromaDB
Specializations: LLM Orchestration, Prompt Engineering, RAG Systems, Agent Architecture, API Development

PROJECTS

Multi-Agent Research Assistant

Built a collaborative multi-agent system using LangChain where specialized agents handle research, summarization, and citation. Reduced research time by 60% for content creation team.

CERTIFICATIONS

- AWS Certified Machine Learning - Specialty
- DeepLearning.AI - LangChain for LLM Application Development