MC

MARIO CUEVAS

mario.cuevas@students.cau.edu | 9145887954 | New York, NY 10027 |
WWW: https://mariostacksnqueues.github.io/Mario-Cuevas-Portfolio/ | **WWW:** www.linkedin.com/in/mario-cuevas-b26421232

Objective

Computer science student with hands-on experience in Python, FastAPI, and Docker orchestration. Successfully architected scalable AI pipelines, reducing server costs by 20% and increasing customer retention by 12%. Seeking internships or full-time roles to contribute to product development and enhance operational efficiency.

Skills

- Python and FastAPI
- Prompt engineering and LangChain
- AWS SageMaker and Lambda
- Docker and Kubernetes
- SQL (PostgreSQL, MySQL) & MongoDB
- Real-time data processing (Spark/Kafka)
- Data Structures and Algorithms
- CI/CD with GitHub Actions

Experience

Automeit.ai | New York, NY
Founder & AI Engineer
12/2024 - Current

- **Processed 5,000 automated tasks per month in a closed beta of our AI agent platform, doubling throughput within two months.**
- Secured three pilot clients—collectively saving them 30 hours of manual work per week—by tailoring agent workflows to their needs.
- **Architected a lightweight, Docker-based orchestration system for LangChain agents, reducing server costs by 20% while improving overall reliability.**

Vedi Robotics | New York City, NY
Software Engineering Intern
05/2025 - 08/2025

- **Developed a Python Flask microservice to serve a trained NLP model via REST API, cutting inference latency by 35%, and enabling real-time sentiment analysis in the company's chatbot.**
- **Implemented a GitHub Actions CI/CD pipeline with Docker for TensorFlow model training and deployment, reducing manual release effort by 75%, and ensuring consistent, reproducible builds.**
- Optimized a vision transformer model on AWS SageMaker through quantization and pruning, shrinking the model size by 60%, and lowering inference costs by 40% per 1,000 requests.

Be.Brooklyn | New York City, NY
Data Science Intern
06/2023 - 08/2024

- Developed an end-to-end data-cleaning and ETL pipeline in Python that reduced raw data processing time by 40% and enabled daily refreshes of our analytics dashboards.
- Engineered a predictive churn-risk model using scikit-learn and XGBoost that achieved 88% accuracy, allowing the marketing team to target at-risk customers, and improve retention by 12%.

Education

Clark Atlanta University | Atlanta, GA
Bachelor of Science in Computer Science
Expected in 05/2026

- Relevant Coursework: Data Structures, Algorithms, Object-Oriented Programming, Software Engineering, Database Systems, Finance, Web Development, Artificial Intelligence

Extracurricular Activities

Student-Athlete (Baseball): Competed at the D1 (Coppin State) and D2 (Clark Atlanta) levels demonstrating elite time-management, discipline, and leadership as a team captain.
Member, AUC NSBE Club: Actively participate in chapter meetings, networking events, and community outreach initiatives.
Tutoring: Mentored 10+ peers weekly in Python, data structures, and algorithms helping raise average student grades by 20%