EUGENE AYONGA

Wilbraham, Massachusetts, USA

+1-585-910-5904 | eugeneayonga@gmail.com | Portfolio: eugeneayonga.com

linkedin.com/in/eugeneayonga | github.com/eugeneayonga

SUMMARY

Founder and Generative AI Data Scientist with an MS in Data Science and a record of building intelligent systems that drive impact across public and commercial sectors. Experienced in designing end-to-end data and AI solutions, from data pipelines on AWS and PySpark to deploying ML, deep learning and LLM models for language, vision, and autonomous agents. Portfolio spans projects and a venture, Vesnay, including large scale sentiment tracking, RAG chatbots, and AI-driven market research tools. Skilled at turning complex analytics into clear, actionable insights, blending technical depth with creativity to build products that learn, adapt, and deliver value.

EDUCATION

University of Rochester **Aug 2023 - Dec 2024**

Master of Science, Data Science *Rochester, NY*

- **Coursework:** Database Systems, Data Mining, Data Science at Scale, Data Structures and Algorithms, Deep Learning, Machine Learning, Natural Language Processing, Probability and Statistics, Time Series Analysis and Forecasting

EXPERIENCE

Vesnay **Nov 2025 - Present**

Founder & CEO *Massachusetts, USA (Remote)*

- Founded VESNAY, an AI wardrobe OS using multi-agent LLMs and concierge workflows to plan outfits by calendar, city, and climate
- Curated product vision, value proposition, and go-to-market for bespoke, luxury wardrobe management for high net worth individuals
- Built the MVP using React Native, Supabase, and agents to ingest wardrobe, weather, and calendar data and generate recommendations
- Built wardrobe graph and recommendation logic to learn each user's rhythm across homes, and seasons to surface context-aware looks
- Identified three prospective pilot users and planned discovery to refine workflows, pricing, and engagement KPIs

City of Rochester, Fire Department **Aug 2024 - Dec 2024**

Data Scientist *Rochester, NY*

- Built pipeline for 1.2M+ incident records using Python, SQL, and PySpark to optimize resource allocation and station planning
- Developed time-series forecasts using Prophet, achieving an 87% accuracy for 10-year demand projections
- Developed interactive geospatial maps with Folium to highlight high-risk zones and support emergency response efficiency
- Led geospatial and temporal analysis and presented executive briefings that potentially drove $800K+ in projected annual savings
- Co-authored research report "Improving Emergency Response Performance at the City of Rochester Fire Department" (2024)

OpenBrand **Jun 2024 - Dec 2024**

Data Science Intern *San Diego, CA*

- Deployed AWS QuickSight dashboards and KPI automation for consumer products, cutting manual reporting time by 70% for analysts
- Built a Machine Learning list price calculator using XGBoost, generating $4,000+ cost savings and improving pricing accuracy
- Presented Business Intelligence insights and visualizations that informed executive decisions and strengthened strategic planning

PROJECTS

Hospital Reviews RAG Chatbot | Project Link **Oct 2025 - Oct 2025**

- Built a retrieval augmented generation system using LangChain, ChromaDB, and OpenAI API to analyze over 10K hospital reviews, achieving sub-2s latency and 15% higher F1/EM through advanced chunking, tuning, and prompt optimization

AI Market Research Agents | Project Link **Sep 2025 - Sep 2025**

- Architected a multi agent workflow combining LangGraph, Serper API, and Playwright to autonomously research and synthesize market briefs, maintaining 92% task success and 95% reliability across diverse analytical domains

Smart Optical Character Recognition Document Automation | Project Link **Aug 2025 - Aug 2025**

- Engineered an OCR document processing pipeline integrating Tesseract, OpenCV, and Gemini API for invoice recognition and validation, automating field extraction and cutting manual review effort by 40%

Tracking Sentiment of Congressional Members' Tweets at Scale | GitHub Link **Apr 2024 - May 2024**

- Built a scalable AWS S3 and PySpark streaming pipeline in Databricks for sentiment analysis of Congressional tweets, deploying a Hugging Face transformer via MLflow (F1 = 0.98) and optimizing Spark UDF inference for high throughput performance

SKILLS

- **Generative AI**: RAG systems, LangChain, Hugging Face Transformers, OpenAI APIs, Vector Databases, Prompt Engineering, LLMs
- **Machine Learning**: Regression, Classification, Clustering, Time series Analysis, Deep Learning, NLP, Computer Vision
- **Data Engineering & Cloud**: Python, R, SQL, PySpark, Databricks, AWS (S3, Bedrock, QuickSight), Docker, MLflow, Git, CI/CD
- **Analytics & Applications**: FastAPI, Tableau, A/B testing, Statistical Modeling, Probability Theory, Data Visualization, Flutter