Blake Martin

blakeatech@gmail.com | linkedin.com/in/blakeatech | www.blakeatech.com | github.com/blakeatech | (270) 570-1840

EXPERIENCE

Data Engineer   Jul. 2024 – Present
CM Incorporated   *Owensboro, Kentucky*

- Designed and deployed a real-time manufacturing planning platform on Azure using FastAPI, PostgreSQL, and Docker; processes 1K+ orders/month, reducing manual planning time by 80% and achieving 99.9% system uptime
- Implemented a custom partitioning engine combining graph partitioning, constraint satisfaction, and k-means clustering to group jobs by compatibility, reducing excess material by 40% and saving $100K+ annually
- Engineered a similarity detection system using weighted Jaccard and cosine distance to group orders by material, size, and processing time, achieving 85% accuracy against historical job pairings
- Developed an async data pipeline with Pydantic-based validation for 20+ derived manufacturing features, including fiber distributions, roll weight variances, and setup optimization metrics
- Built lightweight N8N agents to orchestrate task automation across PostgreSQL, Google Sheets, and email workflows, enabling human-in-the-loop validations and async approvals

Machine Learning Engineer   Jan. 2023 – Apr. 2024
AIC Incorporated   *Palo Alto, California*

- Built an XGBoost model ($R^2 = 0.88$, RMSE = 2.4K) for 3-month sales forecasting using GPT-3.5-turbo-16k for synthetic feature generation, resulting in a 22% increase in quarterly sales
- Deployed model to AWS Sagemaker with Docker; average inference latency: 310ms over 3-month window
- Managed experiment tracking and model versioning with MLflow (v2.5) to ensure reproducibility and streamline model comparison across iterations
- Developed unit tests with pytest and end-to-end tests with Selenium to ensure model reliability and performance
- Deployed asynchronous inference pipelines using AWS Lambda and Step Functions to process webhook-triggered GPT-4 video interview analysis workflows, enabling scalable, event-driven hiring automations
- Integrated a bidirectional LSTM model in PyTorch to classify response fluency and sentiment from transcribed interviews (91.4% accuracy on labeled dataset), enhancing GPT-4 scoring with sequence-aware signal extraction

Machine Learning Intern   Aug. 2022 – Nov. 2022
AIC Incorporated   *Palo Alto, California*

- Built a LangChain agent that evaluated student essay drafts against scholarship criteria and retrieved semantically similar award-winning examples from a corpus of 2000+ essays stored in a Weaviate database hosted on EC2
- Designed the agentic workflow to compare inputs with top-performing essays and generate actionable feedback for improvement, helping 50+ students earn over $120K in scholarships
- Fine-tuned a GPT2LMHeadModel (124M params) on scraped essays to generate proposed rewrites incorporating personalized feedback, deploying through a Gradio interface on HuggingFace

PROJECTS

Notaic | *Python, LlamaIndex, Firebase, Stripe*   github.com/blakeatech/notaic-repo

- Developed an AI-powered email management platform that automates email drafting, achieving over 1,000 user signups and reducing users' email processing time by 70%
- Engineered a personalized knowledge retrieval system using LlamaIndex that processes users' email history to provide contextually relevant suggestions for email responses
- Optimized inference costs by self-hosting Deepseek R1 model on an RTX 4090 via Runpod
- Prototyped multi-agent orchestration using LangGraph to manage distinct agent roles (e.g., retrieval, summarization, and prioritization) for complex email workflows, enabling modular, extensible architecture

EDUCATION

Brescia University   Owensboro, Kentucky
Bachelor of Science in Computer Science   *Aug. 2020 – Dec. 2023*

TECHNICAL SKILLS

Languages: Python, PostgreSQL, Java, Golang, PHP, Scala, JavaScript, TypeScript
Developer Tools: Azure, AWS, Firebase, dbt, Spark, Databricks, Snowflake, Airflow, Docker, Kubernetes, Postman
Libraries: scikit-learn, LlamaIndex, ResNet, YOLO, OpenCV, GPT, LLaMA, BERT, PyTorch, TensorFlow, Pandas