TEMITAYO ABIONA

AI/ML ENGINEER

Email | LinkedIn | GitHub

Tel: +44 77870 19206 Address: Park Royal, London, NW10 7JD.

PROFESSIONAL EXPERIENCE

Machine Learning Engineer (remote): MoniMoore - London, United Kingdom *June 2023 — present*

- Developed a machine learning model for analyzing business KPIs, which improved revenue tracking accuracy by 15% and reduced manual processing time by 90%. These improvements have continued over 18 months, helping the company save £300K annually in operational costs while maintaining a consistent growth rate in revenue tracking accuracy.
- Designed scalable cloud-based infrastructure using AWS SageMaker, optimizing machine learning models for 30% improved performance and 20% cost-efficiency, which was sustained over 12 months, enabling the platform to handle 50% more queries with the infrastructure, ensuring long-term cost savings and scalability.
- Managed LangChain workflows for NLP-based real-time applications, enabling faster query processing and reducing user wait times by 25%.
- Integrated a Retrieval Augmented Generation (RAG) model with Pinecone vector DB, achieving an 87% BertScore. This improved real-time customer query handling and Monimoore's AI-driven income and expense analysis accuracy.
- Designed and implemented scalable cloud-based AI solutions for financial optimization, including fraud detection and real-time income analysis models.
- Led a cross-functional team of 5 engineers and data scientists to integrate AI models into user-facing applications, resulting in a 20% increase in user engagement. This solution has been sustained over two product cycles, leading to a cumulative increase in user engagement by 35% and enabling the team to scale the solution to other products.
- Implemented feedback-driven improvements to AI models, ensuring continuous enhancement of performance. Over 12 months, this approach has led to a 25% reduction in downtime and an increase in system scalability, allowing the model to handle 35% more traffic without additional infrastructure.

AI/ML Engineer: Percent Meta Solutions - Nigeria *Oct 2022 – Jan 2023*

- Conducted model tuning and optimization for scalability and performance improvements, reducing latency by 30%.
- Developed AI-driven solutions integrated with AWS and LangChain, improving automated data collection processes and overall system performance.
- Deployed end-to-end machine learning pipelines using AWS SageMaker, facilitating seamless model deployment and integration with client systems.
- Specialized in prompt engineering and model performance tuning for Large Language Models (LLMs) in real-world business applications.
- Worked closely with clients to prioritize feature requests and ensure that AI solutions aligned with their business goals and operational needs.

EDUCATION

Northumbria University - MSc. Artificial Intelligence Technology, London, United Kingdom

GPA: Distinction

Projects: ***Generative AI for Fraud Detection in PropTech Transactions:*** *Developed a generative AI model that analyzed large datasets to identify fraudulent patterns. This project honed my ability to work with large language models and AI-driven solutions for the financial sector (Achieved Distinction).*

Kwara State University - BSc. Statistics, Kwara State, Nigeria

- *A hybrid dimension reduction technique for prediction of breast carcinomas in patients using information complexity criterion – a grade.*

SKILLS & INTERESTS

Programming Languages: Python, C++, R, JavaScript, React, Stata, Matlab, MERN Stack
Databases & Micro-services: SQL, NoSQL- MongoDB, Cassandra, HBase, DynamoDB, Redis, Docker, Kubernetes
Distributed Systems: Hadoop (HDFS, MapReduce, YARN, Oozie, Hive), ETL (Spark, Flink), Kafka, Storm, BigQuery
ML Platform: PyTorch, TensorFlow, Keras, PySpark, scikit-learn, OpenCV, NLTK, MLflow, Databricks, Airflow, AutoML
AWS: API Gateway, Aurora, Kinesis, S3, CodeCommit, SNS, SQS, EC2, CodePipeline, ElastiCache, DocumentDB, EBS, Monitron, CloudFront, CloudWatch, Translate, Comprehend, Forecast, Lookout, Personalize, Polly, Textract, Rekognition
Machine Learning: ML models & algorithms, Recommender system, A/B testing, Multi-Armed bandit, Ranking models, Probabilistic Models, Classification, Regression, SVM, Naive Bayes, Clustering, Reinforcement Learning, causal inference
Deep Learning: CNN, LSTM, BiLSTM, GANs (DCGAN, CycleGAN, StyleGAN, BigGAN), Diffusion Model (DALL-E 2), CLIP, Transfer Learning, Optimization, Feature Engineering, Hyper-parameter Tuning, Graph Neural Networks, ASR, LLM
Functional AI/ML Skills: Statistics, Probability, Regularization, Feature Engineering, Hyper-parameter Tuning
Pretrained CNN Models: VGG16, VGG19, InceptionV3, ResNet50, InceptionResNetV2, MobileNetV2, EfficientNetB7
Computer Vision: YOLOv3, SSD, Fast RCNN, Faster RCNN, Mask RCNN, UNet, EfficientDet, Vision Transformer (ViT)
Semantic Segmentation Models: UNet, FCN, SegNet, DeepLab, DeepLabv3, Mask RCNN, Panoptic FPN, PSPNet
Python libraries: pandas, numpy, scipy, scikit-learn, HuggingFace, statsmodels, optuna, gensim, spacy, langchain, openAI
Natural Language Processing: Word2Vec, GloVe, fastText; Transformers- BERT, RoBERTa, BART, XLNet, T5, GPT-3
Large Language Models (LLM): GPT-4, Falcon, Claude, Llama2, Alpaca, Vicuna, Mistral, Mixtral8x7B, ChatGPT, Phi
Gen AI Tools: SFT, IFT, PEFT, LORA, QLORA, RAG, RLHF, DPO, Langchain, Vector DB, MoE, Prompt Engineering
Team Collaboration: Led a cross-functional team of 5 engineers and data scientists, fostering a collaborative environment where team members could contribute their expertise. Through regular team meetings, clear task delegation, and open communication, we successfully integrated AI models into user-facing applications, resulting in a 20% increase in user engagement.
Problem-solving: Identified bottlenecks in the machine learning pipeline during a critical project at MoniMoore, where slow inference times delayed real-time customer insights. By implementing advanced optimization techniques and refactoring code, I reduced model inference times by 30%, improving overall system responsiveness and user satisfaction.
Continuous Learning: Learning Computer Vision in depth
Communication: Worked closely with product managers to communicate the impact of new machine learning features in non-technical terms. This helped the team prioritize high-impact features aligned with business goals, resulting in a 20% improvement in project delivery timelines.

CERTIFICATION & PROJECTS

- **AWS Certified Machine Learning – Specialty** (2024)
- **TensorFlow Developer Certification** – in view (2024)
- **Google Professional Machine Learning Engineer**(2024)

UK Work Eligibility: Eligible to work in the UK under a Tier 2 Visa.
Availability: Open to full-time roles, immediately available for remote and hybrid work models.