GRACE UGOCHI NNEJI, PHD

LinkedIn | Google Scholar | ResearchGate | Github

CONTACT

+1-804-701-9614
ugochinneji@gmail.com
Raleigh, North Carolina
Personal Website

EDUCATION

- **Ph.D., Software Engineering** - UESTC | GPA: 4.0 / 4.0 (Top 1%)
- **M.Eng., Software Engineering** - UESTC | GPA: 3.94/ 4.0 (Top 1%)
- **B.Tech., Computer Science** - FUTO | GPA: 3.95/ 5.0 (Top 5%)
- **PGCert, International Education** - Staffordshire University, UK | (Distinction)

CORE SKILLS

- **Languages/Frameworks:**

Python, TensorFlow, Keras, Scikit-learn, Hugging Face, OpenCV, FastAPI, SQL, R, NumPy, Pandas, Scikit-learn, matplotlib

- **Cloud/DevOps:**

AWS, GCP, Azure, Docker, Supabase, GitHub, CI/CD, Terraform, MLflow, Airflow

- **Data/Visualization:**

Power BI, Tableau, Matplotlib, Seaborn, Plotly, Jupyter, Google Colab, Streamlit, Gradio, Excel

- **Expertise:**

Deep Learning, Machine Learning, Generative AI, Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Predictive Modeling, Time Series Forecasting, Multimodal Learning, Explainable AI (SHAP, LIME), Fairness & Bias Mitigation, Feature Engineering, Data Pipelines, Real-Time ML, MLOps, Model Deployment & Monitoring, Anomaly Detection, Edge AI

IMPACT AREAS

Precision Medicine, Medical Imaging, Health Informatics, Energy, Environment, Agriculture, Sustainability Informatics, Smart Systems

CERTIFICATIONS

- TensorFlow Developer Certificate - TensorFlow (2022)
- Machine Learning Specialization - DeepLearning.AI and Stanford University (2023)
- Google Data Analytics Professional Certificate - Google (2024)
- IBM Data Science Professional Certificate - IBM (2024)
- Generative AI: Elevate Your Data Science Career - IBM (2024)
- Developing AI Applications with Python and Flask - IBM (2024)

PROFILE

Lead Machine Learning Engineer with several years of experience designing, deploying, and scaling AI/ML systems in healthcare, energy, and industrial domains. Proven record of delivering production-ready ML pipelines, optimizing deep learning models for real-time performance, and reducing costs through automation. Skilled in LLMs, generative AI, predictive modeling, and cloud-native deployment, with expertise in CI/CD, containerization, and model monitoring. Strong collaborator with experience leading cross-functional teams and driving measurable business impact.

WORK EXPERIENCE

CDUT - Centre for AI Research

Lead AI/ML Engineer March 2025 - September 2025

- Design, develop, and deploy end-to-end ML pipelines (data ingestion, training, evaluation, deployment, monitoring).
- Build and optimize deep learning, LLMs, and generative AI models for production environments.
- Implement CI/CD pipelines for ML deployment, containerization (Docker, Kubernetes), and model serving.
- Ensure models meet performance, scalability, and reliability requirements in real-time applications.
- Collaborate with data engineers and product teams to translate requirements into deployable AI systems.
- Monitor and maintain deployed models, handling model drift, retraining, and performance monitoring.
- Optimize inference speed and computational efficiency on GPUs/TPUs.
- Write clean, maintainable, and well-documented code following software engineering best practices.
- Mentor junior engineers and contribute to code reviews, architecture design, and tech strategy.

Senior AI/ML Researcher August 2022 - March 2025

- Lead applied and theoretical research projects in machine learning, deep learning, LLMs, generative AI, and reinforcement learning.
- Develop novel algorithms and architectures, pushing the state of the art in AI.
- Publish research in top-tier conferences and journals.
- Translate research findings into prototypes and proof-of-concept systems.
- Collaborate with academic institutions, government, and industry partners on AI initiatives.
- Mentor and supervise young scholars, researchers, and engineers.
- Stay ahead of AI trends and evaluate new techniques, frameworks, and benchmarks.
- Write and review grant proposals and secure external funding for research projects.
- Contribute to the organization's long-term AI research roadmap.

UESTC Information Processing Research Institute

ML/DL Research Assistant Engineer Sept. 2019 - July 2022

- Developed a novel model for enhancing medical image quality for the identification of COVID-19 images, achieving an accuracy of 98%.
- Designed innovation model for detecting Pneumonia and diabetic retinopathy using standard medical images.
- Built and deployed an AI-based model for the prediction of breast cancer, achieving predictive accuracy of 97+%
- Mentor and supervise younger staff and team on the technical and presentation skills for international conferences. Team and scholars satisfaction rate at 95%.
- Designed and implemented automated vehicle recognition AI-based system with 97.2% accuracy with automative datasets.

WORK AUTHOURIZATION

- USA PR Green Card Holder
- Open to Remote, Onsite, and Hybrid