Yvonne Okafor

Phone number: (+1) 9125411786 **Email address:** yvokafor@gmail.com

Home: 1400 Statesboro Place Circle, 30458 Statesboro (United States)

ABOUT ME

Graduate student in Computer Science (Machine Learning) with a 4.0 GPA and strong experience translating AI and data-driven solutions into practical, user-focused applications. Skilled in developing structured frameworks, communicating technical concepts to non-technical stakeholders, and supporting cross-functional collaboration. Experienced in designing AI-powered workflows, building analytics tools, and enabling data-informed decision-making. Passionate about responsible AI adoption, employee enablement, and organizational transformation.

EDUCATION AND TRAINING

Masters in Computer Science

Georgia Southern University [06/08/2025 – 05/2027]

Country: United States | **Website:** https://www.georgiasouthern.edu/ | **Final grade:** 4.0

B.Eng Telecommunications Engineering (Graduated Valedictorian)

Bells University of Technology [19/09/2014 – 31/07/2019]

Address: Km. 8 Idiroko Rd Benja village, Ota, Ogun State , Ota (Nigeria) | **Website:** bellsuniversity.edu.ng

WORK EXPERIENCE

Georgia Southern University, IT Services Unit – United States

Graduate Assistant -Computational Research Assistant

[05/08/2025 – Current]

- Designed and implemented automation workflows (using python aumation scipts) that improved accessibility and usability of high-performance computing systems for researchers, reducing manual setup time by approximately 40%
- Developed monitoring dashboards and reporting tools to communicate system usage trends and insights to non-technical stakeholders
- Collaborated with cross-functional teams to support large-scale computational projects, translating technical requirements into actionable workflows
- Assisted users in adopting computational tools by providing guidance, documentation, and troubleshooting support

Outamation Advanced Externship

AI & Data Extraction Extern

[04/12/2025 – Current]

- Designed AI-powered document processing workflows using OCR, NLP, and LLM-based systems to enable efficient information extraction from complex documents
- Built Retrieval-Augmented Generation (RAG) systems to support intelligent search and question-answering, improving accessibility of large document repositories
- Structured end-to-end data workflows and documentation to support usability and adoption of AI tools by diverse users
- Evaluated multiple AI approaches and synthesized findings into clear recommendations for workflow optimization

Gravitas – Lagos, Nigeria

Data Scientist – Retail Analytics & Store Optimization

[02/08/2024 – 31/07/2025]

- Analyzed large-scale transaction data and presented insights through Power BI dashboards and structured reports to support business decision-making
- Led A/B testing initiatives and communicated results using data visualizations and stakeholder presentations, guiding store optimization strategies
- Developed automated analytics workflows and interactive reporting tools using Power BI and Python, reducing reporting turnaround time by approximately 40%
- Translated complex analytical outputs into clear visual narratives and actionable recommendations for non-technical stakeholders

Africartz – Remote

Data Analyst – E-commerce Insights

[14/06/2023 – 10/12/2023]

- Developed Python-based web scraping and API data pipelines using BeautifulSoup to collect, clean, and structure large-scale e-commerce datasets, supporting customer segmentation, churn prediction, and data-driven marketing strategies.
- Conducted exploratory data analysis (EDA) on large-scale e-commerce datasets to identify trends in customer behavior, purchasing patterns, and churn drivers.
- Built customer churn prediction models using Random Forest and Logistic Regression, achieving approximately 85% accuracy in identifying high-risk customers.
- Implemented K-Means clustering and behavioral segmentation to support targeted marketing campaigns, increasing conversion rates by 15%.
- Designed feature engineering and data preprocessing pipelines using Python (Pandas, Scikit-learn) to improve model performance and generalization.
- Developed interactive Power BI dashboards to visualize customer engagement, retention trends, and campaign performance for business stakeholders..

SKILLS

Python / JavaScript / HTML / CSS / PHP / WIX / Wordpress / Research and analysis / Technical Support and Troubleshooting / C++ / SQL / Power BI / Python (Pandas, Scikit-learn) / NLP / RAG Systems

CERTIFICATIONS

[DataCamp]
Understanding Data Science

[DataCamp]
Intermediate Python

[DataCamp]
Introduction to Power BI

[Google]
Google IT Automation with Python Professional Certificate

[Google]
Foundations of digital marketing & e-commerce

[University of Carlifonia, Irvine]
Introduction to the Internet of Things and Embedded Systems

[Coursera]
Use Mailchimp to Build an E-mail Marketing Campaign

[DataCamp]
Supervised Learning with scikit-learn

[DataCamp]
Feature Engineering for Machine Learning in Python

PROJECTS & ACHIEVEMENTS

Projects & Achievements

Privacy-Preserving Healthcare Data System (Project)

- Designed and implemented a hybrid, security-focused ML architecture for distributed healthcare systems, combining Federated Learning to eliminate raw data sharing, Homomorphic Encryption to protect model updates during computation, and Blockchain (Hyperledger Fabric) to provide tamper-resistant coordination, auditability, and trust management across multiple untrusted data owners handling sensitive patient information.
- Developed secure aggregation workflows using TensorFlow Federated and PySyft, reducing gradient leakage attack success rates by ~40% while preserving model accuracy within 2–3% of centralized baselines.

Enhancing Fraud Detection with Machine Learning (Project)

- Developed a fraud detection model using Logistic Regression and Random Forest, achieving a 92% ROC-AUC and reducing false negatives by over 35% compared to baseline models.
- Applied advanced data preprocessing and optimization techniques including data cleaning (pandas/NumPy), feature selection (RFE, SelectKBest), SMOTE oversampling, and robust evaluation (cross-validation, confusion matrix, ROC-AUC) using scikit-learn to significantly improve the predictive performance of fraud detection models.
- Currently developing an interactive, real-time fraud monitoring dashboard using Python and Streamlit/Dash, integrating live model outputs and alerting pipelines to visualize anomalous behavior and support rapid security decision-making.

Bridging the gap between Data Science and Software Development Teams - A Project Management Approach (Research Project)

• Led the design and implementation of a Jira- and GitHub–based project management framework to bridge collaboration between data science and software development teams, ensuring data science work followed Agile planning, GitHub version control, and CI/CD workflows.

• Directed Agile sprint planning, source control workflows, and CI/CD integration, implementing GitHub Actions to enforce testing, code quality, and reproducible development.

• Oversaw debugging and integration of Python-based data pipelines and analytical code, and led design and project management discussions to translate research outputs into deployable software components