

# Peter Ozo-ogueji

+1 (202) 977-1952 | peterchika3254@gmail.com | Washington, DC, USA | LinkedIn | GitHub | Portfolio

## PROFESSIONAL SUMMARY

Results-driven Data Scientist with 7+ years delivering production ML systems that generate measurable business impact. Achieved \$2.1M annual cost savings through fraud detection systems and 30%+ model performance improvements via advanced ensemble methods. Proven expertise in healthcare analytics, financial services, and end-to-end ML pipeline development with strong cross-functional collaboration skills.

## TECHNICAL SKILLS & CORE COMPETENCES

**Programming Languages:** Python, R, SQL, JavaScript  
**ML/AI Frameworks:** TensorFlow, PyTorch, Scikit-learn, Keras, XGBoost, LightGBM  
**MLOps & Deployment:** Docker, AWS SageMaker, Model Deployment, Cross-validation, Hyperparameter Tuning  
**Data Visualization & BI:** Tableau, Power BI, Matplotlib, Seaborn, Plotly, D3.js, Business Intelligence  
**Big Data & Cloud:** Apache Spark, Hadoop, AWS (S3, EC2, SageMaker), Microsoft Azure  
**Databases:** PostgreSQL, MongoDB, MySQL, Redis  
**Statistical Methods:** Regression Analysis, Time Series (ARIMA, GARCH, VAR), Bayesian Methods, A/B Testing  
**Development Tools:** Git, Docker, Streamlit, R Shiny, Jupyter, Google Colab

## EDUCATION

**Dual Degree: MS Business Analytics & AI | MS Data Science** | American University | Washington, DC *Aug 2023 - Dec 2025 | GPA: 3.8/4.0*

**Master of Science, Mathematics** | Philippines College of Health and Sciences | Manila, PH *Aug 2019 - July 2021 | GPA: 93/100*

**Bachelor of Science, Statistics** | Federal University of Technology | Owerri, Nigeria *Sept 2010 - Dec 2015*

## PROFESSIONAL EXPERIENCE

**Data Scientist | METY Technology | Washington, DC** *May 2024 - Present*  
*Healthcare AI startup developing wellness applications for Health and Longevity*

- Developed MyYouthSpan Health App Success Prediction Model using Gradient Boosting and feature engineering  
→ Achieved 84.4% accuracy with cross-validation and hyperparameter optimization
- Presented predictive model insights and ROI analysis to C-suite executives  
→ Directly influenced MyYouthSpan Health App development strategy and feature prioritization
- Conducted comprehensive meta-analysis of 180 health app studies using statistical significance testing  
→ Identified AI-powered features as primary subscription conversion drivers
- Optimized feature combinations through multivariate testing  
→ Achieved 34.7% success improvement with AI + Wearable + Community + Coach combination  
→ Delivered 6 critical marketing factor recommendations

**Graduate Teaching Assistant | American University-CAS | Washington, DC** *Aug 2024 - Present*

- Instructed Statistical Machine Learning and Data Science Applications for 200+ graduate students  
→ Improved course evaluation scores by 20% through innovative Python-based tutorials
- Enhanced student performance in regression analysis, hypothesis testing, and ML concepts  
→ Achieved 20% improvement through hands-on Python tutorials and practical projects
- Developed innovative curriculum in probability theory and statistical decision-making  
→ Increased student understanding by 15% through interactive learning approaches

**Data Analyst | Polaris Bank Limited | Owerri, Nigeria Owerri** *April 2018 - July 2023*  
*Commercial bank with \$2B+ assets serving 500K+ customers across 260+ branches*

- Led implementation of machine learning fraud management system using IBM's big data platform  
→ Reduced fraud incidents by 30% and achieved \$2.1M annual cost savings
- Designed enterprise data warehouse architecture with ETL pipelines and data governance frameworks  
→ Facilitated 35% increase in analytics adoption while reducing infrastructure costs by 15%
- Created 50+ executive dashboards using Power BI and SQL  
→ Supported strategic decision-making with real-time KPI monitoring and predictive insights
- Applied advanced analytics for customer insights using ARIMA time series forecasting, multivariate regression, and clustering

→ Optimized customer segmentation and revenue strategies

KEY PROJECTS & IMPACT

---

Real-Time Auto Loan Pricing Engine - *Link to Project*

Technologies: XGBoost, API Integration, Real-time Data Pipelines

- Built ML-driven dynamic pricing model with real-time rate adjustments and competitive positioning
  - Generated \$5.1M annual revenue improvement (5,150% ROI)
- Achieved 73.47% AUC with processing speeds 2,016x faster than manual methods
  - Integrated live market data from 15+ financial institutions

Customer Churn Prediction Engine - *Link to project*

Technologies: Ensemble Learning, SMOTE, Advanced Feature Selection

- Improved minority class recall from 0% to 99% using advanced SMOTE balancing and ensemble techniques
  - Created \$4.9M annual revenue impact through proactive retention strategies
- Achieved 83.15% AUC enabling targeted customer interventions
  - 95% precision in high-risk customer identification

Speech Emotion Recognition Platform - *Link to project*

Technologies: Python, XGBoost, LightGBM, Random Forest, Streamlit

- Engineered production-ready ensemble ML system with 191 engineered audio features
  - Achieved 82% accuracy across 8 emotion classes
- Developed scalable emotion analytics solution for customer service optimization
  - Potential to reduce manual analysis costs by 70% with real-time insights

Multimodal Hate Speech Detection - *Link to project*

Technologies: BERT, ResNet50, PyTorch, Multimodal Fusion

- Built multimodal ML system combining BERT text analysis and ResNet50 image processing
  - Achieved 66% accuracy on 10,000 memes using advanced fusion techniques
- Engineered high-dimensional feature representations with IVA fusion outperforming CCA by 6% F1-score
  - Created scalable content moderation solution with \$3.5M projected cost savings

LEADERSHIP & ACHIEVEMENTS

---

Google Hackathon Competition Winner | Washington, DC

- Led winning team leveraging SQL, BigQuery, Python, and Tableau to conduct exploratory data analysis
  - Delivered compelling visualizations and solutions for real-world business challenges
- Utilized Google Cloud Platform's advanced analytics tools to tackle complex data problems
  - Demonstrated expertise in data analysis, visualization, and executive-level presentation skills
- Collaborated with multidisciplinary team of professionals showcasing effective teamwork and project management
  - Successfully engaged with Google's Public Sector team and AU Analytics Club stakeholders

Smith Analytics Consortium's 5th Annual Datathon | Washington, DC

- Developed comprehensive NBA expansion strategy using predictive modeling and geospatial analytics
- Selected Seattle, WA and Kansas City, MO as optimal expansion cities through Random Forest algorithms
- Created interactive Tableau visualizations supporting data-driven business recommendations

AI & Robotics Club Executive Member | American University

- Led technical workshops on machine learning implementation and industry best practices
- Increased club membership by 40% through engaging technical content and networking events
- Organized networking events connecting students with data science professionals

CERTIFICATION & PROFESSIONAL DEVELOPMENT

---

**In Progress:** AWS Certified Machine Learning - Specialty

**Completed:** Google Analytics Individual Qualification (IQ)

**Professional Training:** Advanced Statistical Methods, Deep Learning Specialization, Business Intelligence Certificate