Bernard Boateng

Chicago, IL
bbofori90@gmail.com| (312) 960-9527 | LinkedIn

SKILLS

Languages & Frameworks: Python (Pandas, NumPy, Scikit-learn, Flask), R, SQL, NoSQL, JavaScript
Statistical & ML Methods: Statistical Modeling, Machine Learning, LightGBM, XGBoost, Regression, Classification, Time Series Analysis, Root Cause Analysis, Causal Inference, Feature Engineering
Data Engineering & Processing: Data Extraction, Big Data Analytics, Spark, Hadoop, ETL Pipeline Development, Clustered Data Processing
Analytics & Visualization: Data Visualization, Statistical Analysis, A/B Testing, Plotly, Tableau, Matplotlib

WORK EXPERIENCE

Loyola University Chicago **Jan 2025 - Present**
AI/ML Engineer *Chicago, IL*

- Built end-to-end AI-powered Malaria Reprioritization Tool (MRPT) reducing analysis time by 87% (3 hours to 15 minutes), directly impacting bed net distribution for 100,000+ beneficiaries across 15 regions.
- Developed conversational AI interface using OpenAI GPT-4, processing 500+ daily queries with 92% accuracy, enabling non-technical health workers to perform complex geospatial analyses through natural language.
- Architected scalable Flask backend supporting 600 concurrent users, deployed on AWS with 99.9% uptime and sub-2 second response times for complex visualizations.

Loyola University Chicago **Jan 2025 - Present**
Data Scientist *Chicago, IL*

- Deployed LightGBM models predicting high-risk malaria zones with 87% accuracy (5% improvement), potentially preventing $200K in resource misallocation and improving intervention targeting by 30%.
- Conducted causal inference analysis on 15 intervention sites affecting 50,000+ individuals, identifying 3 key factors that reduced transmission rates by 22%, informing $1M+ in funding decisions.
- Built time series forecasting pipeline (XGBoost) reducing utility cost prediction error by 30%, saving $150K annually and improving budget allocation across 20 facilities.
- Automated ETL processes for 10GB+ daily health data from 50 sources, reducing processing time from 6 hours to 45 minutes using PySpark and SQL optimization

Loyola University Chicago **May 2024 - Dec 2024**
Data Science Intern *Chicago, IL*

- Developed ML (Factorization-based Texture Segmentation) model for urban settlement classification achieving 80% accuracy (15% above baseline) on 10,000+ satellite images covering 200 sq km, enabling targeted health intervention planning.
- Processed 5TB+ of satellite imagery and health data using distributed computing, extracting 50+ features for predictive modeling and creating novel urban morphology insights.
- Created automated data pipeline reducing preprocessing time by 70%, enabling weekly model updates instead of monthly, improving response time to emerging health patterns.

PROJECTS

Emergent Intelligence Simulation GitHub

- Developed interactive multi-agent simulation demonstrating collective problem-solving, handling **1,000+ concurrent agents** with real-time D3.js visualization.

Sentiment Analysis Classifier GitHub

- Built and deployed Naive Bayes classifier analyzing **50,000+ movie reviews in under 2 minutes** with **95% accuracy**.
- Production model serving **1,000+ daily predictions**, improving content recommendations and increasing user engagement by **40%**

EDUCATION

Loyola University Chicago *Chicago, IL*
Master of Science in Data Science
Relevant Coursework: Machine Learning, Statistical Modeling, Big Data Analytics, Deep Learning, Data Engineering, Database Management