NANA BOATENG

1115 LANDINGS OVERLOOK• ALPHARETTA, GA 30005 • 901.246.7618

• UNLTD148@YAHOO.CO.UK

SUMMARY

- Nana Boateng is a Data Scientist with more than seven years professional experience in machine learning, Deep learning, natural language processing, optimization techniques, predictive analytics, computer vision, statistical analysis and spatial data visualization.
- Proficient in handling all stages of data science project lifecycle: business understanding, data acquisition & ETL (structured & unstructured data), feature engineering, modeling, deployment and monitoring.
- Built Fraud detection machine learning models for NICE systems Inc. These models are used by major financial institutions around the world to reduce financial crimes.
- Led multiple analytical projects using Customer Usage Data (CUDA) and warranty data to drive insights into customer mileage, identify warranty concerns and improve overall durability of FCA vehicles using analytical tools like R, Python, MATLAB, Tableau and SAS.
- Developed recommender systems for retail marketing based on matrix factorization Techniques such as ALS and Embedding using tools like Spark and Keras.
- Built Machine learning models using libraries such as Scikit-learn, H2O, Caret, Numpy, Matplotlib, Seaborn, Pandas and tidyverse.
- Performed statistical analysis using models such as conditional fixed effects logistic regression for binary categorical outcomes.

Portfolio

- https://nanaakwasiabayieboateng.github.io/
- https://share.streamlit.io/
- https://github.com/NanaAkwasiAbayieBoateng
- https://rpubs.com/mr148

TRAINING AND CERTIFICATIONS

- Introduction to Machine Learning in Production – Coursera, 2022
- SQL and Relational Databases – IBM ,2021
- Certified Base Programmer – SAS
- Data Science with Scala - IBM, 2020
- Docker Essentials – A developer Introduction, IBM, 2020
- American Statistical Association
- American Economic Association
- Institute of Commercial Managers (UK)
- International Conference on Design of Experiments, University of Memphis

TECHNICAL SKILLS

- Programming Languages: Python, R, SQL, Spark, Scala, SAS, Linux/UNIX, Bash, MATLAB, Julia, C++, Octave, Stata, SPSS, Minitab, Datarobot.
- Machine Learning: PyTorch, Tensorflow, Scikit-learn, Tidymodels, Docker
- Cloud: Microsoft Azure, AWS Sagemaker, Databricks
- Version Control: Github, Bitbucket, Gitlab
- Data Visualization: Python (Plotly, Matplotlib, seaborn, bokeh), R(ggplot), Tableau, Power BI.
- Dashboard: Shiny, Dash

PROFESSIONAL EXPERIENCE

NICE Systems Inc.
Senior Data Scientist
May 2019 – present

- Build, test and deploy machine learning models for financial fraud detection with tools such as Datarobot, and Python.
- Monitor and plan resource capacity needs, requirements reporting.
- Manage professional performance, development and mentoring of team members.
- Perform analysis to support the deployment of fraud prevention analytical models.
- Analyze fraud cases obtained from clients.
- Research data patterns to find patterns predictive of fraud.
- Improve the quality and actual implementation of computational algorithms and tools.
- Optimize the detection performance of NICE Actimize Fraud products and improve customers' experience with our Fraud solutions.
- Define product requirements for analytics and provide feedback to the product team on ways in which product may be improved.
- Develop and enhance our solution-specific risk scores.
- Measure the quality of the analytical performance of Fraud Products.
- Develop tools to support model tuning, performance tracking and automation.
- Develop custom detection logic for specific clients.
- Develop tools for data drift detection and model monitoring.

AWARDS

- NICE Sparkhaton winner 2021- Deep learning for fraud detection project
- NICE Impacts awards 2022 winner - fraud detection models developed for a major financial institution saved them nearly 17 million annually.

Oglethorpe University
Part -Time Faculty of Mathematics
January 2021 – December 2022

- Instructor for undergraduate mathematics and statistics. Taught Introduction to Statistics (Spring and Fall Semesters)

Georgia Gwinnett College
Part -Time Faculty of Mathematics
January 2020 – December 2020

- Instructor for undergraduate mathematics. Taught Precalculus.

The R Journal
Reviewer (Volunteer)
January 2019 - present

- Reviewer for several R packages for the premier R language software journal

Catalina Marketing Corporation, GA
Senior Data Scientist
December 2018 – March 2019

- Built recommender systems for CVC/SharebuildR campaigns with Matrix factorization methods such as ALS and Embedding with tools such as Spark and Keras. Explored Distributed Learning tools such as Elephas, Dist-Keras and Horovod Runner.

- Built an end –to-end flow model to rank propensity and redemption forecast models with tools such as Scikit-learn, Spark.
- Build visualizations and dashboard using shiny to display forecast from a revenue management-forecasting model.
- Mentoring Associate Data Scientist.

Fiat Chrysler Automobiles, MI
Data Scientist
June 2017 – December 2018

- Pothole and significant events detection with machine learning. Used Machine learning models including MLP, XGBoost and AutoML to detect and to predict pothole size.
- Lead manpower requirements project to predict with better accuracy the number of vehicles that would be sent to Chrysler Proving Grounds for vehicle testing. This allowed the manpower team to hire the right number of drivers thereby reducing cost otherwise incurred from hiring more drivers than will be needed.
- Sentiment Analysis of FCA employee and ex-employee reviews: Scraped and analyzed thousands of employee and ex-employees reviews from Glassdoor and Indeed between 2008 and 2018.

- Lead on multiple analytical projects using Customer Usage Data (CUDA) and warranty data to drive insights into customer mileage, identify warranty concerns and improve overall durability of FCA vehicles.

- Lead Trailer Tow project using data from Control Tec database to analyze 95th Percentile trailer towing FCA SUV vehicles.
- Member of team developing Qlikview interface to various FCA vehicle databases Participated in weekly meetings to analyze various stages of the development Qlikview.

Environment: R, Python, Spark, MATLAB, Tableau, SAS, SQL, Qlikview

Baptist Cancer Center, TN
Data Analyst / Manager
January 2016 - May 2017(Concurrent with Jobs Below)

- Responsible for data management that includes data collection and database management for the Thoracic Oncology Multidisciplinary Clinic.
- Duties include collecting data at conferences and during clinics and reporting to Medical Director, Medical Steering Committee, Administration and various grant-funding organizations as directed.
- Perform a prospective matched cohort comparative effectiveness study of patients receiving serial versus multidisciplinary care, with key patient-centered endpoints (survival, satisfaction with the care experience, timeliness and appropriateness of care, quality of staging).
- Perform statistical analysis to determine the quality of care and survival between multidisciplinary program and serial care program using models such as conditional fixed effects logistic regression for binary categorical outcomes; fixed effects generalized linear models and fixed effects proportional hazard model for survival analysis.

Baptist College of Health Sciences, TN
Adjunct Instructor
August 2015 - August 2016 (Concurrent with Job Below)

- Taught College Algebra and Introduction to Statistics.

The University of Memphis, TN
Graduate Assistant
August 2012 - August 2016 (Concurrent with Job Above)

Member of Professional Development Committee, January 2014 - August 2016

- Recruited resource personnel for graduate student professional development seminars.
- Organized professional development seminars.

Member of Graduate Assistants Healthcare Committee, January 2014 - August 2016

- Partnered with Church Health Center to bring health insurance coverage to Graduate Assistants.

Graduate Teaching Assistant, August 2012 - May 2016

- Taught Introduction to Statistics, Foundations of Math, and Elementary Calculus.

St. Jude Children's Research Hospital, TN
Graduate Research Assistant
July 2014 - June 2015 (Concurrent with Jobs Above)

- Worked on numerous research projects, including:
- (1) Gastronomy complications in pediatric cancer patients.
- (2) The effect of dynamic contrast: enhanced MRI(DEMRI) on tumor angiogenic activity and in predicting tumor response of OS2008 patients.
- (3) Evaluation of alternative in vivo screening methodology: Analysis of single mouse tumor response results from PPTP.
- (4) The significance of splenomegaly at diagnosis in pediatric Hodgkin lymphoma.

Middle Tennessee State University, TN
Graduate Teaching Assistant
August 2010 - August 2012

- Taught College Algebra.
- Tutored Calculus I, II &III, Partial and Ordinary Differential Equations, Financial Mathematics and Trigonometry.

Publications

- Paul Azunre, Salomey Osei, Salomey Addo, Lawrence Asamoah Adu-Gyamfi, Stephen Moore, Bernard Adabankah, Bernard Opoku, Clara Asare-Nyarko, Samuel Nyarko, Cynthia Amoaba, Esther Dansoa Appiah, Felix Akwerh, Richard Nii Lante Lawson, Joel Budu, Emmanuel Debrah, **Nana Boateng**, Wisdom Ofori, Edwin Buabeng-Munkoh, Franklin Adjei, Isaac Kojo Essel Ampomah, Joseph Otoo, Reindorf Borkor, Standylove Birago Mensah, Lucien Mensah, Mark Amoako Marcel, Anokye Acheampong Amponsah, James Ben Hayfron-Acquah: English-Twi Parallel Corpus for Machine Translation, arXiv preprint arXiv:2103.15625, 2021

- Fernandez Israel, John A. Sandoval, Reagan M. Jones, **Nana Boateng**, Jianrong Wu, Bhaskar N. Rao,. Andrew M. Davidoff, Stephen A. Stochat. Gastrostomy Complications in Pediatric Cancer Patients: A Retrospective Single-Institution Review, Pediatric Blood & Cancer 62(7):S184-S185, November 2015.

- **Nana Boateng**, Numerical Partial Differential Solution of The Black-Scholes Equation, Lambert Academic.Publishing (2013).

- Brendan Murphy, Han Yin, John Maris, E Kolb, Richard Gorlick, C. Patrick Reynolds, Min Kang, Stephen Keir, Raushan Kurmasheva, Igor Dvorchik, Jianrong Wu, Catherine Billups, **Nana Boateng**, Malcolm Smith, and Peter Houghton. Evaluation of Alternative In Vivo Drug Screening Methodology: single mouse analysis: A Retrospective Single-Institution Review, Cancer Research, October 2016.

Technical Experience

- Bayesian Inference (A/B Testing).
- Deep Learning/Distributed Deep Learning tools, Spark ML, Keras, Tensorflow, Elephas, Dist-Keras, Horovod .
- Multi level/Hierarchical Modeling.
- Natural Language Processing with Large Language Models

- Predictive Analytics and Time Series Modeling tools scikit-learn, H2O(R and Python).
- Survival Analysis.
- Unsupervised learning: Clustering Algorithms, Dimensionality Reduction (PCA, t-SNE, UMAP).
- Machine learning (deep learning, recommender systems, natural language processing, network analysis).
- Non-parametric Regression and Methods.
- Diagnostic Imaging and Sensitivity Analysis.
- Computational Modeling and Biostatistics in Gene Expression Data.
- Sentiment Analysis (structured and unstructured data sets).
- Sample Size Estimation and Power Analysis.
- Analysis of Count Data.
- Spark (SparkR / Sparklyr / PySpark)
- Optimization techniques (linear / nonlinear programming, dynamic / stochastic programming).
- Quantitative finance (Monte Carlo simulation, risk quantification, portfolio optimization, economic scenario generation).
- Parallel, GPU and Cloud computing(Databricks, Google Cloud Platform, Elastic cloud in Azure).
- Spatial Data Visualization.

EDUCATION

- PhD, Statistics (2016), The University of Memphis, Memphis, TN
- Master of Arts in Economics (2017), The University of Memphis, Memphis, TN
- Master of Science in Statistics (2014), The University of Memphis, Memphis, TN
- Master of Science in Mathematics (2012), Middle Tennessee State University, Murfreesboro, TN
- Graduate Diploma in Business Studies (2010), Institute of Commercial Managers, United Kingdom
- Bachelor of Science in Mathematics (2007), Kwame Nkrumah University of Science and Technology, Kumasi-Ghana