Kingsley Gaa
Senior Data Scientist

Albany, NY - Email me on Indeed: indeed.com/r/Kingsley-Gaa/d887c9b1fd430dd8

Willing to relocate: Anywhere
Authorized to work in the US for any employer

WORK EXPERIENCE

Senior Data Scientist
Sekaf Ghana Ltd - November 2016 to April 2017

Tamale, Ghana
• Developed Java application code contained in JAR files and deployed it to SAS to enable SAS to connect to Hadoop
• Connected SAS to Hadoop cluster, submitted a request and get it processed for the results to be sent back to SAS
• Imported data into R for data exploration and data cleaning for developing predictive models as per requirements
• Implemented large scale data and analytics using advanced statistical and machine learning models to classify similar products together for developing sub-markets

Data Analyst / Graduate Student
University at Albany, SUNY - Albany, NY - August 2014 to August 2016

• Developed a project on structural models to produce Multivariate time series forecast of consumption, investment, interest rates and other endogenous stochastic variables in SAS
• Built Logistic Regression Models to predict Probability of Default (PD)
• Created a project for predicting future sales volumes with SAS skills, including SAS/MACRO, SAS /IML, SAS/GRAPH, SAS/BASE, and SAS/STAT
• Developed Two Stage Least Square models to solve system of equations for predicting future loan demand, interest rates, etc. and performed Monte Carlo simulation using SAS

Senior Data Scientist
Sekaf Ghana Ltd - March 2009 to July 2014

Tamale, Ghana
• Used PROC SQOOP to access Apache Sqoop utility from a SAS session and to transfer data between a database and HDFS
• Implemented Machine learning algorithms on big data to find natural patterns in data that generate insight into customers' purchasing behavior and help make better decisions and predictions
• Analyzed data sets to help transition credit risk computations, programming in R, python, and Java / Hadoop to deliver solutions
• Built demand forecasting models to predict the trend and future volumes of import/export products utilizing Time Series ARIMA programming in SAS

Statistical Analyst
Ford Motor Company - Dearborn, MI - April 2007 to January 2009

• Worked in large scale database environment like Hadoop and MapReduce, with working mechanism of Hadoop clusters, nodes and Hadoop Distributed File System (HDFS)

• Applied Python and machine learning algorithms such as Logistic Regression, SVM, KNN, K-means and Random Forest to developed insight into customers' purchasing behavior and help make better decisions and predictions
• Created structural models to produce Multivariate time series forecast of retail sales and performed Monte Carlo Simulation for evaluating forecast performance in SAS software
• Predictive modeling experience in using logistic regression to predict binary outcomes for credit risk analysis with the use of SAS
• Experience in SAS programming of simultaneous equations to forecast economic conditions on interest rate, real GDP, inflation, and investment etc.
• Created SAS macros and developed Generalized Linear Models (GLM) using nonlinear regression techniques to predict Benchmark Auction prices for off-lease vehicles
• Used CART, CHAID, decision trees, and Data Mining Techniques in SAS to develop classification models
• Built predictive models to forecast residual values of leased vehicles and estimated the demand for these used vehicles using Regression and Time Series techniques in SAS
• Pulled data from Oracle database using SQL queries, data analysis, generating reports, graphs and corrected data step syntax errors

Statistician / Data Analyst

Accra, GH - October 2000 to August 2005

Ghana

• Collected data on economic indicators, analyzed economic trends, and developed models to forecast growth, interest rates, inflation and investments using SAS
• Programming experience with SQL and SAS, including relational database query construction, optimization and predictive modeling
• Practical experience in employing multivariate regression techniques to identify trends in market demand and consumption
• Developed segmentation of customer base by utilizing Multivariate Statistical Techniques, such as Clustering and Factor Analysis using SAS and Excel to get potential target customers

EDUCATION

Master of Arts degree in Economics

State University of New York, University at Albany   -   Albany, NY
May 2017

Graduate Certificate

State University of New York, University at Albany   -   Albany, NY
May 2017

Master of Arts in Economics

Miami University, Richard T. Farmer School of Business, Oxford
December 2006

Bachelor of Arts in Economics

University of Ghana
June 1998

SKILLS

• Developed a project on structural models to produce Multivariate time series forecast of consumption, investment, interest rates and other endogenous stochastic variables in SAS • Built Logistic Regression Models to predict Probability of Default (PD) • Created a project for predicting future sales volumes with SAS skills, including SAS/MACRO, SAS /IML, SAS/GRAPH, SAS/BASE, and SAS/STAT • Developed Two Stage Least Square models to solve system of equations for predicting future loan demand, interest rates, etc. and performed Monte Carlo simulation using SAS

ADDITIONAL INFORMATION

• 10+ years of experience of Data Modeling and Analytics in creating SAS programs to produce ARIMA forecasts and made prediction of stochastic endogenous variables
• 5+ years of experience working as Data Scientist in applying machine learning and statistical modeling techniques to solve business problems
• Connected SAS to Hadoop cluster, extracted data, and delivered it to SAS for processing and for developing predictive models
• Manipulated, merged and created restructured datasets from big data using Machine learning, Hadoop, SAS, Python, R, SQL and built predictive models
• Experience in data mining, including predictive behavior analysis, CHAID, CART, Optimization and Customer Segmentation analysis using SAS and SQL
• Possessed outstanding communication skills and ability to handle multiple projects within a given timeframe and work cooperatively in a team

COMPUTER AND QUANTITATIVE SKILLS

• Machine Learning • Hadoop / Big Data • Java programming
• Python • C# programming • SQL programming and Oracle
• Statistical Modeling • Regression Modeling • Data mining
• Predictive Modeling • Time Series Forecasting • R programming
• SAS programming • Multivariate Statistics • MapReduce
• Excel, Word, PowerPoint • Access • Visio
• Tableau • Business Intelligence • Knowledge in HTML and CSS
• Data Visualization • Minitab • Applied Mathematics