**SEAN RIGGS**

(980) 241-7331 Charlotte, NC dsriggs1@gmail.com

|  |  |
| --- | --- |
| **ANALYTICAL AREAS OF EXPERIENCE** | |
| Hadoop Distributed File System | Data Engineering |
| ETL  Automation | Shell Scripting  Version Control |
| Relational Databases | Data Visualization  Unix |

Analytical Software: SAS, Python, R, SQL, PySpark, Tableau, GitHub

MS Skills: Excel, PowerPoint, Word

**EDUCATION**

**University of North Carolina Charlotte, Graduated with M.S. Economics (2014-2016)**

* Awarded merit based graduate assistantship
* Graduate Econometrics, Advanced Business Forecasting, Advanced Macroeconomics, Financial Econometrics, Financial Management

**North Carolina State University, Graduated with B.S. Economics (2009-2014)**

* Graduated Cum Laude

**WORK EXPERIENCE**

**Wells Fargo-Quantitative Associate** (June 2018-Present)

* Constructed data mart off transactional level monthly deposits data containing billions of records with over 20 years of history. Applied data mining techniques to categorize raw transaction statement descriptions into different transaction indicators.
* Migrated data mart process from SAS/SQL to python based big-data platform using PySpark, resulting in a 50% reduction in processing time. Further enhanced process by using shell scripts to run the ETL process for multiple months in parallel.
* Partnered with external MIT/IBM team and built deposit account dataset with 20 years of account balance and geographical data. Served as point of contact to assist in understanding the data so that they could test machine learning models to forecast balances.
* Linked internal Wholesale accounts to BEA consumer spending by NAICSUBSECTOR to help inform management of risk of exceeding asset cap. Used Tableau visualizations to highlight estimated impact to Wells Fargo balances if consumer spending exceeded pre-COVID19 levels.
* Built datasets from transactional level data and developed features to aid model developers in identifying accounts in troubled positions and the risks of aggregate losses from those accounts. Developed dashboard visualizations from transactional data using Tableau and regularly presented findings to management and relevant stakeholders.

**Bank of America-Quantitative Finance Analyst** (August 2017-April 2018**)**

* Responsible for running as many as 20 statistical tests as part of the validation process for logistic regression credit scorecard models. Key responsibilities include modifying and developing SAS Macros to perform key statistical tests to evaluate model accuracy, discriminatory power, and sensitivity to changes in model parameters.
* Worked with developers to understand complex methodologies and data manipulations such as the creation and replication of pseudo default datasets used for scorecard modeling.
* Developed challenger models with alternative inputs and data manipulations to provide effective challenge to models submitted by developers.
* Performed quarterly ongoing monitoring for 10 credit scorecard models, and documented results using Latex for typesetting.

**Wells Fargo-Forecast Analyst /Analytic Consultant** (September 2015-August 2017)

* Used SAS Macro language programming to quickly loop through multiple forecasting models to efficiently back-test alternative predictive models. Used both multiple regression analysis, and Box-Jenkins time series analysis to select the best model. Used automated code to back-test challenger models using cross-validation, and holdout sample. Presented findings to management and business partners.
* Re-developed Bankruptcy inflow forecasting model using multiple regression model with seasonal adjustment that resulted in forecasting error being reduced by more than 50% for both short and long-term forecasts.
* Took initiative to integrate R functionality within the SAS environment through Proc IML. Educated forecast team on the capabilities of using R and SAS together, and lead effort to automate forecasts using user-built R functions.
* Worked with other teams to employ use of Box-Jenkins Methodology to identify seasonality in time series and select appropriate ARIMA forecasting model specification.
* Automated forecasting models and KPI metrics using both SAS language, as well as advanced excel VLOOKUP and match index functions.
* Developed Service Release forecasting process and expanded it from three line of businesses to encompass all of default servicing. Communicated regularly with forecast owners for each line of business.
* Leveraged SQL server database to automate manual reporting tasks that had previously been done in Excel by building forecast history SQL table to automatically update KPI accuracy metrics. Developed complex SQL queries using subqueries to pull data from multiple data sources, and perform data transformations.
* Developed 10 ad hoc forecasts across the Bankruptcy business in support of capacity tool development to help senior leaders, and business partners to better understand the key drivers of the Bankruptcy forecast.
* Responsible for tracking forecasting accuracy across multiple lines of business, and using these accuracy metrics to determine where improvements in forecasting methodology can be made. Developed KPI metrics to track accuracy using various metrics, and time intervals.

**PERSONAL PROJECT**

**Fantasy Baseball: Developed machine learning model for estimating fantasy baseball performance based on historical Major League Baseball data** (January 2017-Present)

* Analysis in R using dataset with over 12 million observations and over 200 columns with data from 1952-2017. Analysis focuses on data from 2005 to the present due to fundamental changes in the game that impact offensive performance, and to exclude bulk of steroid era that greatly inflated offensive player performance.
* Data munging using base R aggregate function to transform raw data from plate appearance level to game level to reduce noise.
* Game level data split into four matchup datasets of batters and pitchers based on dominant hand to take advantage of baseball splits.
* Data mining was used for feature engineering of modeling inputs. Used combination of R functions: Data.table, and rollapply to create rolling averages of player performance. Used for loop to create averages of statistics from 1-10 games. Created statistical player performance inputs based on rolling time periods and batter vs pitcher matchups. Input features consist of statistical, categorical, and weather inputs.
* Predictive modeling goal is to predict the probability that player scores some non-zero amount of fantasy points and based on ranking of probabilities, player performance is mapped back to discrete point counts.
* Predictive modeling approach is a stacked ensemble approach combining the results of individual machine learning algorithms: Logistic Regression, Random Forest, Gradient Boosting, and Neural Network. Algorithms are estimated and then combined through the H2o package in R. Optimal lineup is then selected based on predicted points subject to the salary cap constraint.
* Created functions in R for diagnostic testing, testing includes: measuring discriminatory power of model for rank ordering using modified Somers’ D statistic, analysis to quantify for incorrectly ordered pairs how many notches off the prediction is from reality, and the correlation in the predicted point counts with actual counts.