**SEAN RIGGS**

(980) 241-7331 Charlotte, NC dsriggs@ncsu.edu

|  |  |
| --- | --- |
| **ANALYTICAL AREAS OF EXPERIENCE** | |
| Regression Analysis/Modeling | Back Testing |
| Statistics | Machine Learning |
| Data Analysis | Data Cleaning/Manipulation |
| Time Series Forecasting | Predictive Modeling |

Analytical Software: SAS, R, SQL, STATA, MATLAB

MS Skills: Excel, PowerPoint, Word

**EDUCATION**

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

* Awarded merit based graduate assistantship
* Coursework: 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**

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

* Responsible for running as many as 20 statistical tests as part of the validation process for logistic regression credit scorecard models. Key components of this task include modifying and developing SAS Macros to perform key statistical tests which 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 an 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 led effort to automate forecasts using user-built R functions.
* Worked with other teams to employ Box-Jenkins Methodology to identify seasonality in time series; and select appropriate ARIMA forecasting model specifications.
* Automated forecasting models and KPI metrics using SAS language as well as advanced excel VLOOKUP and match index functions.
* Developed Service Release forecasting process and expanded it from three line of business 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 planning 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 used these accuracy metrics to determine where improvements in forecasting methodology could 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 the bulk of the 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 batter/pitcher pairs based on dominant hand to take advantage of baseball splits.
* Data mining was used for feature engineering of modeling inputs. Used combination of the 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/pitcher matchups. Input features consist of statistical, categorical, and weather inputs.
* Predictive modeling goal is to predict the number of fantasy points that a player will score in a given game. The approach taken is first predicting the probability that player scores a non-zero amount of fantasy points; and then based on a ranking of these probabilities, player performance is mapped back to discrete point counts.
* The predictive modeling approach is a stacked ensemble approach combining the results of individual machine learning algorithms: Logistic Regression, Random Forest, Gradient Boosting, and Neural Networks. 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, which 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.
* First generation of model to be implemented for the 2018 season. Further analysis for the future will include improvement for modeling unknown matchups of batters vs relief pitchers, and developing further diagnostic testing for individual algorithms, ensembled model, and full system modeling approach. Potential weighting will be explored for giving more weight to data from recent years over older years.