RYAN JONES

DATA SCIENTIST

CONTACT

407–924–4588
rybojones.github.io
ryan.wesley.jones@gmail.com

EDUCATION

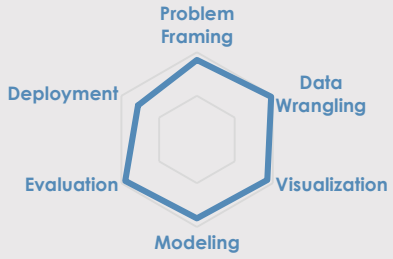
MSc Data Analytics
University of Central Florida, May 2021
Distinction: Exceptional Graduate Student Award in Data Analytics (4.0 GPA)

BSc Mathematics + Minor Statistics
University of Central Florida, May 2015

CERTIFICATIONS + CLEARANCE

Public Trust Clearance
AWS Cloud Practitioner
AWS Machine Learning Specialty *(Feb 2022)*
Udacity Nanodegree - Programming for Data Science

CRISP-DM SKILLS



TOOLS + KNOWLEDGE

Programming: Python, R, SQL, Git, SAS, Bash, C, Java,JavaScript, HTML, CSS, Agile Development

Analysis: AWS, PowerBI, Tableau, PostgreSQL, MS SQL Server, Snowflake, Excel, Observable

Machine Learning Libraries: Pandas, Numpy, Scikit-Learn, PyCaret, MLflow, XGboost, Pyspark MLLib, Seaborn, Plotly, Matplotlib, NLTK, Spacy, Gensim, Keras, TensorFlow, Tidyverse, GGPlot

Machine Learning Algorithms: Linear Regression (OLS, Ridge, Lasso, ElasticNet, PLS, PCA Regression), Non-Linear Regression (Linear Methods w/ Transforms, Support Vector Machine, K Neighbors Regressor), Classification (Logistic Regression, Decision Tree, Random Forest, LightGBM, XGBoost, K Nearest Neighbor, Naive Bayes), Bootstrapping, Bagging, Boosting, Clustering (K Means, Hierarchical, DBSCAN, Agglomerative), Time Series Analysis (Moving Average, Auto Regression, ARIMA), Neural Networks (RNN, LSTM, Transformer, Auto Encoder, GAN, CNN), Feature Importance (Permutation, Shapley, Tree Based)

Other: CRISP-DM, Instructional Design, Content Delivery and Presenting, Consulting Best Practices, Project Management

SNAPSHOT

Data scientist with experience and ambition, driven by a desire to deliver impactful and responsible models. I bring passion to projects and strive to fully understand processes to identify the best solution to difficult problems. I continually devote considerable time and effort into practicing, refining, and learning the required skills to successfully exercise robust data science principles in my endeavors, both professional and personal.

EMPLOYMENT

Deloitte Consulting - Data Scientist, USDC
June 2021 - Present

- Development of regression, classification, and time series models for customer experience measurement, cost forecasting and marketing focus.
- Text analytics of customer verbatim for deep customer experience analysis.
- Quality assurance and validation of statistical methods.
- Strategic advisement of data architecture for customer journey mapping and advanced analytics.

Adapt Health - Data Analyst, Advanced Analytics Group
August 2019 - May 2021

- Development of unsupervised learning model for insights into network structure and traffic.
- Agile software development of internal CRM tool (product owner).

Wekiva High School - Mathematics & Statistics Teacher
June 2016 - July 2019

- PLC leader; content and delivery design.
- Remediation planning through performance analysis.

PROJECTS

Network Graph of Call Center Traffic
Developed a novel network graph solution to visualize customer call traffic patterns across thousands of nodes (customer service agents). This technology helped to improve overall customer experience (+21% customer ratings) by enabling management to reroute calls to avoid bottlenecks and remove unwanted loops that kept callers from reaching a final destination.

Time-Series Modeling for Heavy Equipment Lifetime Cost Prediction
Utilizing SCRUM methodology, delivered a time-series regression model predicting the point in time that the maintenance cost per hour became greater than the purchase cost per hour for heavy equipment. This allowed the organization to increase customer retention and sales through targeted marketing campaigns.

Regression Modeling with Driver Analysis for Customer Experience
Generated customer overall satisfaction regression models by combining survey and operational data. Identified key drivers, and corresponding importance, to guide resource allocation decisions as part of the broad goal of customer experience improvement.

Binary Classification for Electric Vehicle Charging Station Mapping
Working with large data (25+ billion records) in a Snowflake data warehouse, disaggregated electric utility load usage to classify customers with EV charging stations. Defined a novel classification model that outperformed traditional machine learning models by achieving, on average, ~10%+ higher accuracy, precision and recall. This enabled the organization to identify geographical areas to target for investment.

Binary Classification Using Machine Learning Ensemble
Collaborating with a team of close female peers in the Women in Data Science 2021(WiDS) worldwide data-thon organized by Kaggle, we created an ensemble model utilizing XGBoost and Linear Regression to identify ICU patients with Diabetes Mellitus. We were able to achieve an AUC score a mere 0.010 below the winning submission score.

Collaborative Filter Recommendation Engine Utilizing Parallelized Computing
Constructed user-movie ratings models, utilizing user-defined KMeans and Alternating Least Squares algorithms, and with PySpark as part of a graduate school project focusing on machine learning using cloud computing techniques to identify consumer propensity towards unseen movies.