Ryan Jones

(407)924-4588 / ryan.wesley.jones@gmail.com / rybojones.github.io

I have always been intrigued by mathematical and statistical analysis, understanding the processes involved in real-world phenomena, and identifying solutions to difficult problems. In the past, this natural curiosity steered me to study numerous topics while in school and later helped me to graduate with a bachelor's degree in mathematics and a minor in statistics. On top of this, I have a *profound* interest in data analysis and modeling. My appetite for learning has led me to embark on a continuous quest for knowledge. I devote considerable time and effort into learning, refining, and practicing the required skills to successfully employ data science principles throughout my career.

Education

M.S. Data Analytics
University of Central Florida, Graduation date of May 2021

Curriculum focusing on statistical methods, preprocessing data, visualization, machine learning, cloud computing and data storage techniques. All classes culminate in a project designed to test knowledge of learned concepts.

B.S. Mathematics + Minor Statistics
University of Central Florida, May 2015

Applied mathematics and statistics curriculum included machine learning adjacent studies, including differential equations, linear algebra, and numerical methods.

Work Experience

Business Analyst *August 2019 – present*
AeroCare Holdings Inc. (now AdaptHealth)

Managed migration of high-value report data structures into a new Azure instance. Performed documentation and provided high-level summarization of SQL batch jobs to leadership in an effort to consolidate reports across two companies all while preserving the reporting requirements set forth by the business SMEs.

Math & Statistics Teacher, *August 2016 – June 2019*
Wekiva High School

Increased Probability & Statistics student test scores by 6%, on average, over the year prior by using targeted remediation derived from data analysis.

Named *Most Impactful Teacher* by top 5% graduating senior.

Completed Projects

Binary Classification Using Domain Knowledge

Winner of UCF Big Data Analytics Symposium 2021 competition with the stated goal of disaggregating Orlando Utilities Commission customer load usage and classifying customers with an electric vehicle. Working with large data (25+ billion records in some tables) in a SnowFlake data warehouse, we were able to define a domain knowledge model that outperformed traditional machine learning models by achieving, on average, 10%+ higher accuracy, precision and recall.

Consumer Propensity Model

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

More information at https://github.com/rybojones/MSDA-Recommendation-System

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, placing us firmly in the top-half of submissions. This experience was both exciting and educational and I plan on participating in future Kaggle events frequently to get exposure to new methodologies and continually grow my knowledgebase.

Natural Language Processing and Contextualization

Utilized natural language processing toolkits to perform sentiment analysis and topic modeling for all U.S. Senate and House of Representative member Tweets in the month leading up to the 2020 Presidential election. These findings provided insight into the limitations of sentiment analysis on social media text and how context can be obfuscated when using such a technique.

Visuals at https://public.tableau.com/profile/ryan.jones1301#!/

Open Projects

Time-Series Modeling and Prediction

Working as a member of a small consultancy team for Southern States ToyotaLift, I am tasked with predicting the point in time that the maintenance cost per hour becomes greater than the purchase cost per hour for a forklift. My team is using the SCRUM software development process and GitHub version control tools as we define the best model to fulfill the defined business need.

Optical Character Recognition (Upcoming)

With friends and members of my MSDA 2021 cohort, we intend to create a free-for-use tool that identifies text in an image, using OCR, and subsequently translates that text to the Braille print file format. We feel that this would help to reduce persistent costs surrounding replicating print material for the visually impaired, especially in underfunded special needs classrooms.

Tags

Python, SQL,R, Bash, Excel, Tableau, PowerBI, Git, Pandas, Numpy, SciKit-Learn, NLTK, Keras, PySpark, MPI4Py, Cloud Computing, Machine Learning, Matplotlib, Seaborn, Plotly, C, JavaScript, RMarkdown, Shiny, GGPlot, Tidyverse, Jupyter, Observable, Jira

Visit my personal website at rybojones.github.io to learn a little more about me!