

JOHNATHON SMITH

Data Scientist



713-203-3603



smith.johnathon01@gmail.com



<https://bit.ly/30bBTqj>



<https://bit.ly/3n9ltHH>

Data Scientist and US Navy veteran experienced with the practical application of data analysis, machine learning, and forecasting. Tenacious problem solver with a knack for creative solutions. Attacks large, complicated tasks with a focus on teamwork and skillful collaboration. Dependable professional desiring to be a useful member of the team and provide value to the business as a whole.

Inactive Secret Clearance, lapsed Feb 2018.

TECHNICAL SKILLS

Python - MySQL - Tableau - Git - NumPy - Pandas - Scikit-learn - Matplotlib - Seaborn - Jupyter - Excel - Applied Statistics - SciPy - Machine Learning - Natural Language Processing - Time Series Analysis - Anomaly Detection - Data Storytelling - PySpark - Web Scraping

WORK EXPERIENCE

H-E-B, Refrigeration Maintenance Tech

May 2019 - Dec 2019

Recorded plant data and used it to identify malfunctioning components on a daily basis. Clearly communicated my findings with leadership, resulting in more efficient maintenance periods and fewer equipment failures.

US Navy, Nuclear Machinist

Feb 2012 - Feb 2018

Analyzed both historical and real-time data to identify long term trends and anomalous conditions in the plant. Used domain knowledge of ship operations and plant conditions to guide watch stander action with data-driven decisions. Quickly identified and communicated the information needed to recover from various casualty situations, minimizing plant down time and preventing loss of ship's power.

EDUCATION

Codeup

December 2021

Certificate of Completion

Fully-immersive, project-based 22-week career accelerator that provides students with 670+ hours of expert instruction in applied data science. Students develop expertise across the full data science pipeline (planning, acquisition, preparation, exploration, modeling, delivery), and become comfortable working with data to deliver actionable insights to diverse stakeholders.

University of Texas - San Antonio

April 2021

Computer Science, 87 Credit Hours

GPA: 3.52

DEVELOPMENT PROJECTS

Esports Analysis, Dec 2021

Classification

Esports is a booming global industry that will soon rival that of traditional professional sports. League of Legends is one of the largest esports in the world that generated \$1.75 billion dollars for Riot Games in 2020 alone. A professional match can last around 20 minutes, and we've created a model that predicts the winner of a match at the 10 minute mark. Our best Random Forest Classifier model performed 23% better than baseline and had an accuracy of 61% on unseen test data.

Predicting Github Programming Languages, Nov 2021

Natural Language Processing

Using K-Nearest Neighbors, Decision Tree, Random Forest Classifier, Gaussian N-Bayes, and Multinomial N-Bayes models, we predicted a Github repository's primary programming language based on the text content of the README file. Ultimately, we built a model that could predict whether or not a repository was primarily Javascript with an accuracy of about 95% on test data.

Austin Animal Center: Intakes and Outcomes, Oct 2021

Time Series Analysis

Acquired data sets from Austin's Open Data Portal and merged them into a single data frame on a common datetime index for time series analysis. Utilized simple average, moving average, Holt's Linear Trend, and previous cycle models to predict animal intakes. My best model (Holt's Linear Trend) had an RMSE of 471 animal intakes on test data.

Predicting Property Values, Sep 2021

Regression

Utilized LinearRegression, LassoLars, TweedieRegressor, and Polynomial models to predict the tax value of California properties. My best model (Polynomial, degree = 3) had an RMSE of \$194,000 and R² value of 34% on test data.