


ANTHONY RIVERA-STRAINE

DATA SCIENTIST

 github.com/datastraine/

 /anthony-straine

 anthony.straine@gmail.com

 317-529-9798

CORE SKILLS

- Python for Data Science
- Data Analysis
- Intelligence Analysis
- Data Management
- Policy Implementation
- Training and Development
- SQL
- Business Intelligence
- Cognos Analytics
- Knowledge Management
- Talent Management

ABOUT ME

I am an innovative Data Scientist and BI analyst with the ability to translate raw data into easy-to-understand information while providing actionable solutions to business problems. As an intelligence analyst with the Army, I developed a passion for data analytics and finding the needle in the haystack. I have a track record of accomplishing measurable results while leading teams of up to ten people in a dynamic, fast-paced environment. I've conducted analysis and production of various intelligence and business products focused on Business Intelligence, Predictive Analysis, Knowledge Management, Data Visualization, and Counterintelligence. My career is supported by two Bachelor's Degrees.

PROJECTS

Credit Risk Analysis (Current complete 2021)

Chinese banks are very risk-averse, yet younger people are taking on more debt. What are the contributing factors to defaulting on credit card debt? How can we know if an account will default in the future? That is what my team aims to find out! We will accomplish this by looking for patterns in the user's credit history and application data, with a watchful eye to ethical consideration, and then build a model to predict whether an account will default six months into the future.

Predict Facebook Ad Category

Utilizing data gathered by the AdObserver project, predict whether a Facebook ad is political or not political. The end result was a logistic regression model that had an F1 Score on political ads of 90%

Analyzing Git Readmes with NLP

After scraping the readme and primary coding language text from over 200 GitHub Repos utilizing BeautifulSoup, my team used the results from Term Frequency – Inverse Document Frequency (TFIDF) of the readmes to train a logistic regression model in order to predict the coding language used by a project.

Zillow Clustering Project

The goal was to identify key features and clusters, created using KMeans, that had the highest impact on "log error". The impact on "log error" was determined by running a logistic regression using the features, while the clusters were tested for statistical significance utilizing Chi-Squared tests.

Zillow Price Prediction

My team set out to predict the sale price of a property utilizing both given and engineered features. After determining which features correlated with the sales price, we built a Lasso Lar's polynomial prediction model to predict the price.

WORK EXPERIENCE

DATA SCIENTIST

IBM | Nov 2018 - July 2020

- Conducted an in-depth analysis of services' free/trial tier to determine the business viability of each free/trial tier and general usage metrics.
- Created a comprehensive Structured Query Language (SQL) code library to increase the efficiency and reach of reports.
- Created and implemented team-wide Cognos Framework Manager and Cognos Analytics protocols and standards.
- Built various Cognos Analytics business intelligence reports used by top executives to track business metrics such as monthly recurring revenue, average revenue per account, and customer churn.
- Performed real-time, ad-hoc analysis to answer business and data analytic questions such as product life cycle and revenue, while translating data into layman terms.

SIGNAL INTELLIGENCE ANALYST

U.S. Army | Sep. 2011 to Oct. 2018

- Conducted general pattern and correlation analysis and in-depth calculation analysis of large data sets using SQL and National Security Agency (NSA) databases to create accurate products.
- Created clear and concise intelligence reports to provide situational and strategic awareness of operations to customers; provided real-time intelligence processing and briefings to leadership and customers.
- Managed a team of 5 intelligence analysts assigned to find, analyze, and disseminate information pertaining to threats in several areas of interest in global locations.
- Worked across various organizations to draft and finalize signal intelligence concepts for several exercises and real-time operations.

SECURITY CLEARANCE

TS-SCI | LAPSED OCT 2020

SCHOOLS ATTENDED

CODEUP | JULY 2020- JAN 2021

Data Science

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 real, messy data to deliver actionable insights to diverse stakeholders

NORWICH UNIVERSITY | AUG 2004 - MAY 2008

Bachelor of Arts in Political Science

Bachelor of Arts in Criminal Justice

- Corp of Cadets Graduate