

MARLEY ROBINSON

Data Scientist

 (210) 744-0655  MarleyCRobinson99@gmail.com

 /Marley-C-Robinson-99

 /marley-c-robinson

Data Scientist and Engineer graduate from a 22-week immersive coding bootcamp with intensive knowledge and experience in the data science pipeline such as stages including modeling, data acquisition and preparation, and the implementation of data-driven solutions. Technical professional utilizing previous educational history in linguistics to facilitate and develop computer language-based skillsets as an engineer and programmer.

TECHNICAL SKILLS

Python | SQL | Jupyter | Deep Learning | Bioinformatics | VSCode | Data Science Pipeline | Data Science Libraries | Classification | Regression | Clustering | Natural Language Processing | Web Scraping | Regular Expressions | Anomaly Detection | Time Series | Machine Learning Models | Applied Statistics

EDUCATION

Codeup Data Science Program Jul 2021 - Dec 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.

Applied Statistics - SQL - Python - Pandas - Matplotlib - Seaborn - Plotly - Machine Learning - Natural Language Processing - Apache Spark - Data Storytelling - Git - Jupyter Notebooks - Anaconda - Tableau

University of Texas at San Antonio Aug 2017 - Dec 2020
Bachelor of Arts in Modern Language Studies

Northwest Vista College Aug 2016 - July 2017
Associate of Arts

WORK EXPERIENCE

Texas Landscape Care Nov 2020 - Current
Freelancer/Site lead

Worked on a team to provide landscape services to customers in the greater San Antonio area. As a site lead gained leadership experience in delegating tasks, time management, establishment and maintenance of workplace culture, managerial duties, and training new employees.

The Garden Center Dec 2018 - Aug 2019
Sales Associate

Worked as a cash controller, diagnosed plant illnesses and recommended curative or preventative solutions, maintained item and inventory information in the local inventory database, operated and maintained heavy machinery such as bobcat skid loaders and electric carts, and assisted team members with various tasks often requiring two or more people to complete.

Additionally, led and created content for classes aimed at educating visitors in the basics of plant care, lawn maintenance, and hydroponic cultivation methods.

PROJECTS

Drug Discovery

Regression/Classification/Bioinformatics/Team Project
Predicting the effectiveness of chemical compounds for treating a target disease. This prediction can be used to pre-screen drugs before expensive and time-consuming lab experiments.

My role in the team is that of a machine learning engineer, with my primary focus being to work to provide accurate models utilizing robust machine learning techniques such as grid search, K-fold cross-validation, and deep learning techniques such as the creation of artificial neural networks.

Github Programming Language Prediction

NLP/Classification/Team Project
Employed NLP language processing techniques to clean, stem, lemmatize, and vectorize over 5,000 README files acquired by web scraping Github repos and performed classification to predict the primary programming language of a repo.

Worked specifically in the modeling stage of the pipelining to optimize model performance to achieve a 94.8% prediction accuracy on out-of-sample test data.

Histone 4 Occupancy

NLP/Classification/Bioinformatics
Parsed sequences of genetic information using NLP techniques such as 'Bag of Words' count vectorization, K-merization, and tokenization to classify each sequence of 500 nucleotide bases into a positive or negative class for H4 histone nucleosome occupancy. Created Artificial Neural Networks among other algorithms to predict the class of each sequence and achieved an 84.4% accuracy with a Multinomial Naïve Bayes classifier.

Predicting Zestimate Log Error

Regression/Clustering
Utilized various regression models and clustering models in an effort to find features that directly contributed to a decrease in the log error of Zillow's Zestimate using a 2017 dataset of Zillow single-unit properties. After using a Linear Regression model, I outperformed my baseline RMSE by ~.02 points.

Estimating Home Values

Regression
Predicted the values of single unit properties using a variety of features. Focused on delivering accurate predictions for properties sold during the "hot months" of real estate sales. My best performing Linear Regression model had an R-squared value of 0.33 and an RMSE of \$270,211.