

SUMMARY

Skilled Data Analytics professional with expertise in Python, SQL, R, and Tableau, I possess extensive experience in managing the data science pipeline. My capabilities span from data wrangling and exploratory data analysis to data visualization, hyperparameter tuning, and feature engineering. Additionally, I have a proven track record of applying machine learning models to big data with significant results. Currently, I am seeking a challenging Data Scientist position where I can leverage my analytical abilities and proficiency in driving business success.

TECHNICAL SKILLS

Programming: Python, SQL, R, Java, MATLAB, HTML, CSS, Bash, Scala

Libraries: NLTK, Scikit-learn, Pandas, NumPy, dplyr, H2O Flow, PySpark

Data Collection: JSON, CSV, API, Web-Scraping

Data Visualization: Matplotlib, Seaborn, Plotly, ggplot2, Power BI, Tableau

Database and Cloud: AWS (S3, EC2, RDS), Git, ETL, Database Management (MySQL, Oracle SQL, PostgreSQL), Dashboarding

Probability and Statistics: Statistical inference, Hypothesis testing, Experimental design, A/B Testing, Bayes Theorem, Aggregation

Predictive Modeling: Classification, Regression, Linear Regression, Logistic Regression, Random Forest, Naive Bayes, Clustering, Machine Learning, Supervised, Unsupervised

PROJECTS

Text Classification, Sentiment Analysis of COVID-19 Tweets using Natural Language Processing 2022 - Current

- **Tools:** Python, NLTK, Matplotlib, Plotly, Seaborn, Scikit-learn, Pandas, Numpy
- Developed NLP-based supervised machine learning models for Text Classification and Sentiment Analysis on Covid-19 Twitter data, uncovering valuable insights and driving informed decision-making
- Gathered structured data from Kaggle, performed tokenization, lemmatization, TF-IDF to assist with feature engineering and improve model performance
- Model the training data to several linear (Logistic Regression, Random Forest, Naive Bayes, Support Vector Machine, and non-linear models (Extreme Gradient Boost), achieving **64%** accuracy

AirBnB New York City Price Prediction using Machine Learning 2022 - 2023

- **Tools:** Python, Plotly, Seaborn, Matplotlib, Pandas, NumPy
- Design and implementation of a supervised machine learning model in python on a real-world Kaggle dataset, featuring a comprehensive data science pipeline to predict prices in five NYC neighborhood groups
- Scraped publicly available structured dataset from Kaggle, removed outliers and data discrepancies, handle missing values, scaled categorical features, pre-processing and feature selection
- Modeling the training dataset to multiple machine learning models (linear regression, decision tree, bayesian regression, ridge, and lasso regression, XGBoost), achieved an **R-squared = 57%, MSE = 0.19, & MAE = 0.14**

EXPERIENCE

Springboard, *Data Science Fellow*, Houston, TX 2022 - Current

- Proficient in Python and SQL, with a strong background in exploratory data analysis, data visualization, hypothesis testing, and machine learning
- Completing over 600 hours of the extensive data science curriculum, demonstrated expertise through the completion of three in-depth capstone projects: (github.com/ttariqaziz)

Senor Enterprises Inc, *Data Analyst*, Houston, TX 2019 - 2022

- Evaluated, developed, and maintained financial reports to clearly communicate actual results, implemented more rigorous check and balance policies to avoid transaction mistakes
- Utilized advanced data analysis skills to visualize, mine, and manipulate a 5-years database using Excel and Tableau resulting in a 10% increase in branch performance and revenue with an estimated 80k annual savings

Robert Half Technologies, *Service Engineer-1*, Fremont, CA 2018 - 2019

- Furnished technical and warranty support at level-1, conducting hardware/software diagnostics for a range of systems including servers, blades, and PCs
- Coordinated, oversaw, managed, and documented faults using the ticket system, ensuring prompt resolution of issues and timely completion of projects

EDUCATION

Sofia University, CA

Masters in Computer Science (minor: data science)

FAST-National University, International

Bachelors in Telecommunication Engineering (minor: computer science)