

Charlie Tran

713-835-7531 charlietran01@gmail.com <https://www.linkedin.com/in/charlietran01> <https://github.com/charlietran01>

EDUCATION

University of Houston, Houston, TX

Bachelor of Science, Mathematics – Data Science

December 2024

Minor: Data Science

GPA: 3.6

Relevant Coursework: Statistics for Political Science, Probability, Data Science and Statistical Learning, Machine Learning, Data Programming, Data Science I & II

San Jacinto College, Pasadena, TX

December 2022

Associate of Science, General Studies

GPA: 4.0

Relevant Coursework: Advanced Mathematics - Differential Equations

SKILLS AND CERTIFICATIONS

- **Programming Languages:** Python, C#, R, Java, SQL
- **Software:** Microsoft Office, Excel, Tableau, Google Drives, Unity
- **Libraries/Frameworks:** Python- Pandas, Tensorflow, Scikit-learn; R- ggplot2, tidyverse, Knitr, e1071, caret
- **Microsoft Certifications:** Azure Fundamentals DP-900, Power Platform PL-400

EXPERIENCE

Optimizing Well Site Selection for Hydraulic Fracturing with Deep Learning, University of Houston December 2024

- Built machine learning models (Random Forest, RNN, LSTM, SVM) to predict optimal well locations, achieving Normalized **MSE of 0.0038** and **R² of 0.9999** for production predictions.
- Engineered features from oil/gas production and seismic activity data, integrating geospatial analysis using **GeoPandas** and **Contextily** for mapping insights.
- Conducted hyperparameter tuning (**GridSearchCV**) and applied regularization to improve model performance, mitigating overfitting risks.

Predicting Term Deposit Subscriptions Using Machine Learning, University of Houston

November 2024

- Developed and compared machine learning models (SVM and KNN) to predict term deposit subscriptions using the Bank Marketing dataset, achieving **accuracy of 87.76%** and **F1-score of 92.32%** with SVM.
- Preprocessed data using frequency encoding, imputation, and feature scaling to address class imbalance and enhance model performance.
- Analyzed decision boundaries and model interpretability, highlighting SVM's robustness in high-dimensional data.

Real Estate Pricing Analysis in NYC, University of Houston

December 2023

- Conducted **linear regression analysis** in **R** to evaluate the relationship between property size and price.
- Created visualizations (scatterplots, histograms) using **Tableau** to identify trends and data variability.
- Found an extremely weak correlation (-0.003) with high price variability, concluding no significant relationship (p-value = 0.874, R² = 0.00001).

WORK EXPERIENCE

Data Specialist, Ameritas Insurance

October 2022 – June 2023

- Processed Dental Claims from clients and providers and helped the claims processing team increase productivity and meet goals with individual accuracy and efficiency
- Used statistical methods to calculate the variance that would predict the number of claims needed to be processed per day to meet goals for payment to clients

ITS Help Desk – Student Worker, Macalester College

September 2020 – May 2022

- Communicated at the front desk by replying to students', professors', and faculties' emails/tickets regarding technical issues related to the college's software/hardware.
- Promoted to the back desk in September 2021 to work on imaging laptops for students and professors to use.
- Supported full-time staff with setting up new faculty computers and emergency tech issues during classes.