Xiaoyu Fu

(408)-921-2158 | xiaoyufu1997@gmail.com | Linkedin: www.linkedin.com/in/xiaoyu1997 | GitHub: https://github.com/LarryF97

SUMMARY

Recent graduate of a data science certification with a strong foundation in statistical modeling, machine learning, and data manipulation. Skilled in Python, R, and scikit-learn for developing predictive models and extracting insights from complex datasets. Adept at presenting complex findings to technical and non-technical stakeholders. Eager to leverage skills to drive data-driven solutions for business growth and informed decision-making.

SKILLS

LANGUAGES: Python (Pandas, NumPy, Scikit-Learn, PySpark), SQL
DATA COLLECTION: JSON, CSV, API, Web-Scraping
DATA VISUALIZATION: MatplotLib, Seaborn, Plotly, Tableau, Power BI
MACHINE LEARNING: Hypothesis Testing, A/B Testing, Linear/Logistic Regression, Classification, kNN, Random Forest, Naive Bayes, K-Means Clustering

RELATED PROFESSIONAL EXPERIENCE

Springboard-Data Scientist Trainee   December *2022 – September 2023*

- Participated in 500+ hours of hands-on course material, with 1:1 industry expert mentor oversight.
- Completed 3 in-depth portfolio projects.
- Mastered skills in Python, SQL, data analysis, data visualization, hypothesis testing, and machine learning.

BCG-Data Science & Analytics Virtual Experience   *May 2020 – June 2020*

- Cleaned and checked out the data to address missing values, duplicates, and data type conversions, statistics and correct missing data.
- Built a churn model and predicted the probability of customer churn by code script to enhance customer screening criteria.
- Created and developed an executive summary to let customers get the right data and visual experience.

PROJECTS

Capstone #1: British Airways Passenger Booking   *June 2023*

- Objective: Created a trained model and developed on the provided dataset to predict consumers' sales channels to book airline vacations.
- Tools: Python, Pandas, Seaborn, NumPy, SciPy, LightGBM, ROC-AUC
- Outcomes: Based on speculation, the completion rate increases in direct proportion to the number of services and booking channels from 10.68% for 0 services to 18.59% for 3 services.

Capstone #2: Big Mountain Resort   *March 2023*

- Objective: Created a pricing model that can determine a competitive price to customers, and reflects the advantages of the facilities of the big mountain resort.
- Tools: Python, Pandas, Seaborn, NumPy, SciPy, XGBoost, LightGBM
- Outcomes: Big Mountain Resort's modeled price is $92.96, and the actual price is $81.00 so even with the expected mean absolute error of $10.42, this suggests there is room for an increase.

EDUCATION & CERTIFICATIONS

Springboard Data Science Certificate   *September 2023*

6-month intensive course in data science, machine learning, Python, and SQL with 1:1 industry mentor oversight.

San Jose State University   *May 2023*
Bachelor of Science: *Mathematics*

Introduction to Abstract Mathematics and Proofs, Vector Calculus, Modern Geometry and TransformationsAbstract Algebra, Ord. Diff. Equations and Dynamical Systems, Applied Probability, and Statistics.