

# Jiaxin “Tammy” Tang

tammytongcn@gmail.com | 209-305-9802 | [Linkedin](#) | [Github](#)

## Summary

A full-stack data engineer with robust STEM background and 4+ years of sufficient project experience in programming, data analysis and machine learning. Involves in **building data ETL pipeline, data cleaning, feature engineering, data visualization, predictive modeling and analytical reporting** with big data warehouse. Has a wide variety of research experience in e-commerce, education and finance industry. Passionate about data mining, self-motivated and able to collaborate in a diverse team environment.

## Relevant Experience

### Data Engineer (Contract)

12/2023 – present

@Bookiea LLC | Cupertino, CA

- Designed and implemented robust and scalable database schemas to support B2B and B2C business goals.
- Developed comprehensive sets of key performance indicators and metrics to assess and monitor database performance, ensuring optimal functionality and reliability.
- Worked closely with internal stakeholders and the Stripe team to develop strategy for long term data platform architecture.
- Integrated Stripe payment platform based on our business goals using JavaScript and PostgreSQL.
- Built an algorithm in Python to predict fraudulent transactions and prototyped an alert system to prevent fraudulent activities.
- Applied exploratory data analysis on 138K+ transactions, preprocessed data by removing duplicates, encoding categorical features and handling imbalanced labeled data by SMOTE.
- Trained supervised machine learning models including Logistic Regression, Random Forest and XGBoost via k-fold cross validation and found optimal model parameters via grid search.
- Selected the best model based on confusion matrix/AUC and designed a real-time alert system to prevent fraudulent activities.

### Senior Associate – Business Analytics

04/2019 – 08/2023

@Dream Builder Financial | San Jose, CA

- Conducted extensive marketing research and satisfied each client’s individual needs by providing products/services based on customer-centric and data-driven use cases, professional in risk mitigation and asset management.
- Applied data mining techniques to various financial models from different financial providers to align each client’s financial objectives, risk tolerance, return on investment, developed overall personalized financial plan and devised presentation deliverables.
- Led a team of 10 to conduct marketing campaigns, contributed over 50% of the copywriting for the external advertisements, improved metrics gathering, drove rebranding strategy for 80% of the product offerings by collecting, analyzing and presenting data with MS Office Suite/GSuite.
- Provided trainings to the team on fraud detection and anti-money laundering, stay up-to-date on the latest fraud prevention strategies and techniques as well as suspected fraudulent activities reporting procedures in accordance with state/federal laws and regulations.
- Collaborated with case manager, data analyst and underwriter to bring analytical rigor and statistical methods to monitor clients’ long-term and short-term investment goals, as well as developed innovative solutions with marketing, engineering and UX personnel through agile framework meetings.

## Education and Certificates

**Master of Science in Engineering/Data Science** @UC Riverside | Riverside, CA

12/2023

Related Courses: Machine Learning, Data Mining Techniques, Systems Engineering, Info Retrieval, Database Management Systems

**GPA: 3.91/4**

**Database and Data Analytics Certificate** @UC Santa Cruz Extension

06/2021

Related Courses: Data Analysis with R, Big Data Tools and Use Cases, Internet of Things, Big Data Applications, DB Design and SQL Programming

**Programming in Python** @De Anza College | Cupertino, CA

06/2021

Related Courses: Python Programming, Advance Python Programming, Java Programming

**GPA: 3.5/4**

**Bachelor of Science in Science Education/Physics** @Guangxi Normal University | Guilin, China

07/2013

Related Courses: Probability and Statistics, Linear Algebra, Calculus, Physics Experiments, Literature Retrieval, C++ Programming

**GPA: 3.54/4**

Publication: Jiaxin Tang, Zizhen Zhao, *A New Determination of the Fusion Heat of Ice*, Journal of Experimental Science and Technology, 2012

## Skills

**Programming:** Python (Sk-Learn, Pandas, Numpy, SciPy, Beautifulsoup, TensorFlow, Pytorch), R (dplyr, pastecs, rcompanion), Spark, Java, Linux

**Visualization:** Power BI, Tableau, MATLAB, Matplotlib, Seaborn, ggplot2, Advanced Excel, PowerPoint, Adobe Creative Suite, GSuite

**Database management:** ManageBac, AWS, Hadoop, Google Cloud, Databricks, MySQL, PostgreSQL, NoSQL, MongoDB, Kafka

**Machine Learning Techniques:**

- Linear Regression, Logistic Regression, Support Vector Machine, Decision Tree, Random Forest, K-Nearest Neighbors, Principal Component Analysis (PCA), K-Means Clustering, Ensemble Learning (Random Forest, AdaBoost, Gradient Boosting, XGBoost), Embeddings
- Exploratory Data Analysis, Regularization, Feature Engineering, Model Evaluation, Experimental Design, A/B Testing