

Renfei Gao

1952 Harrison St, Santa Clara, CA, 95050
215-470-6192, gaorenfeifei@gmail.com

U.S. Permanent Resident

6 years of data analyst experiences, manipulating and digging data in **SAS, SQL** and **Python**. Experience in modeling development using machine learning method such as **Random Forest, Xgboost, Logistic Regression, Naïve Bayes** and **SVM**.

SKILLS

- **Programming Languages:** Python, SQL, R, SAS, Matlab.
- **Machine Learning:** Regression, Statistical tests, Data Visualization, Random Forest, Deep Learning, Clustering, Naïve Bayes.

DATA SCIENCE PROJECTS

- **Predict Lending Club loans default rate to identify Frauds.** **January-March 2018**
 - Researched and created creative features, like some ratios (credit_to_annuity, credit_to_goods), ltv and etc. to get more hidden information and get the model improved.
 - Performed various sampling methods on highly imbalanced data to increase the prediction by 10%.
 - Grid Search reasonable model parameters to reduce the overfitting and improve the testing score for 10%.
 - Cleaning customer loan data, impute missing value and create new features for final data.
 - Visualized loan data to get more insights for the data.
 - Tested on multiple classification methods like Random Forest, Logistic regression, and Gradient Boosting to make the model performance better.
 - Choose reasonable threshold to make the true positive rate better and get more frauds detected.
- **Pricing Strategies** **May-June 2018**
 - Cleaning data, one-hot encoding the category variables, impute missing value and create new features.
 - Visualized pricing data to get more insights for the conversion rate.
 - A/B testing the different price strategy to choose the better one to get significant revenue improvement.
 - Building model to get the solution for improving the conversion rate.
 - Ensemble multiple models like Random Forest, boosting and Naïve Bayes to make the model performance better.
- **Predicting Salary** **May-August 2018**
 - Building NLP model to predict salaries.
 - Analysis salary using job title, job description and other information related to salary.

EXPERIENCE

Quantitative Analyst

Bank of the West
March 2014 – Present

- Worked on CCAR Modeling and Analytics.
- Build model to predict the default ratio and loss, to help the bank evaluate the risk for the next year.
- Create the module to build the database for portfolios in the bank, including Auto, Mortgage, Whole Sale, Dealer Flooring, Credit Card and etc.
- Automatic ETL code to build modeling data and make it fast with macro code. Improve the process from 2 days to 4 hours.
- Researched and implemented regularized regression techniques for variable selection.

Data Analyst

US Food and Drug Administration
February 2012 – March 2014

- Data manipulation and dataset preparation: importing, cleansing, merging, combining, splitting, and validating.
- Prepare CDISC data set for drug efficient modeling development.
- Dealt with over a million of data records with approximately 400 variables and perform **statistical analysis**.
- Prepared detailed SAP (**Statistics Analysis Plans**) from study protocol with statisticians.
- Advanced experience in many areas of statistics, including linear single and multiple-variable regressions, logistic regression categorical data analysis, experimental designs, etc.

CERTIFICATION AND EDUCATION

- **Data Science Program(Springboard)** **March 2018-June 2018**
Highlights: **Python, Sklearn, Random Forest, Naive Bayes, Logistic Regression, Linear Regression, KNN, SVM**
- **M.S. in Financial Engineering,**
Temple University, Fox School of Business, Philadelphia, PA **August 2010-June 2011**
- **B.S. in Applied Mathematics**
Wuhan University, School of Mathematics and Statistics, Wuhan, China **September 2006-June 2010**