Boyi Chen

Los Angeles, CA | joecby@gmail.com | +1 949-572-3603

EDUCATION

University of California – Los Angeles, CA (QS 29) **June 2022 - September 2025**

Statistic

Major Courses: Statistical Programming with R, Programming in Python, Programming in C++, Linear algebra, Probability, Statistical Methods, Numerical Methods, Economics.

University of California – Irvine, CA **June 2020 - September 2022**

WORK EXPERIENCE

Suzhou Qianmu information technology Co., Ltd, China **June 2021– Aug 2021**

Intern Software Development Engineer

- Provided technical assistance by maintaining the backend MySQL database. Additionally, I leveraged Django, Nginx, and uWSGI to construct a web interface that facilitates the analysis and processing of database information. This interface efficiently presents the data in both tabular and graphical formats, enhancing user accessibility and interaction with data.
- To ensure the effectiveness of the new software for managing raw wood storage, I conducted field inspections at our storage facilities to understand the specific needs and usage habits of the workers. Feedback indicated that the software's data entry interface was overly complex and non-intuitive. Collaborating with my team, we refined the interface based on this feedback, resulting in a more user-friendly solution that significantly enhanced the workers' efficiency in using the software.

Shiseido Company, Limited, China **July 2023 – September 2023**

Intern Data Analysis

- Identified high employee turnover as a critical challenge; executed comprehensive data analyses using R to uncover key drivers of attrition, and formulated strategic interventions, achieving a 16% reduction in turnover rates.
- To support the company's decision-making for next year's sales volumes and budget allocations for various departments, I proactively communicated with financial managers across departments, gathering market data from the past five years. By employing regression analysis, I developed a detailed forecasting report, which was later reviewed by the company's Chief Financial Officer (CFO), providing solid data support for the company's budget planning and strategic decisions.
- Faced with the cumbersome task of manually updating Excel-based HR reports monthly, I identified and resolved issues with data uniformity and gaps by cleaning and reorganizing the data to establish a more efficient local database. Subsequently, I employed PowerBI to visualize the HR data, creating an interactive report that centrally displays key information.

Deloitte Touche Tohmatsu CPA Co.,Ltd. Beijing Branch **July 2024 – September 2024**

Intern Data Analysis

- Partnered with a leading bank to develop a scoring model for evaluating corporate loan default risks, addressing challenges in assessing risk across diverse industries and company sizes. The project required analyzing five years of historical loan data while collaborating with internal stakeholders to meet specific needs.
- Assisted in mapping workflows to identify unique scoring standards for 17 industry categories and various company sizes, including large enterprises, SMEs, and micro-businesses.
- Supported the development of a dual-model system: Quantitative Model: Engineered metrics such as ROA, ROE, and EBIT margins using 62 raw variables, customized for five business categories. Optimized variables using genetic algorithms, PD curve analysis, and other advanced techniques to ensure stability and predictive accuracy. Qualitative Model: Developed tailored variables through surveys distributed to branch managers, collecting expert insights on factors like sales stability and regional growth potential. Selected key variables using Somers' D and industry expert reviews.

PROJECTS

Predicting Alcoholic Status Using Health Vitals **Winter 2023**

- Collaborated on a team project to develop a predictive model for alcoholism using data from the National Health Insurance Service in Korea.
- Led a data analysis project to enhance model accuracy, starting with a General Linear Model and progressing through multiple advanced frameworks. Transitioned from XGBoost to LightGBM due to its efficiency in handling complex data interactions, achieving significant performance improvements. Employed rigorous parameter tuning and iterative refinements, this iterative approach demonstrated my ability to adapt and innovate within high-stakes, competitive environments.
- Competed in a Kaggle challenge to evaluate our predictive model's accuracy in forecasting alcoholism risk, using key health indicators. Our model scored 0.73306, ranking us third globally and first among peer submissions, highlighting our data-driven decision-making skills and our model's robustness.

ADDITIONAL INFORMATION

Languages: Mandarin (native), English (fluent).

Programming: Python, Java, C/C++, R, MATLAB.