BINNY TSAI

ENGINEERING & DATA SCIENCE

CONTACT

- ✉ binny.tsai@gmail.com
- ☎ (626) 417-8309
- 🪪 Chino Hills, CA
- in linkedin.com/in/binny-tsai
- ⭘ github.com/tracy15932

EDUCATION

2017
UNIVERSITY OF CALIFORNIA, LOS ANGELES
B.S. in Mathematics/Economics with Minor in Statistics

TECHNICAL SKILLS

- Python, R, SQL (MySQL)
- Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn
- Apache Spark
- Html/CSS/JavaScript
- Linux, Bash
- AWS (EC2, S3)
- Machine Learning
- JIRA, Git
- Tableau
- VS code, Spyder, Jupyter Notebook

OTHER SKILLS

- Languages: English, Mandarin

EXPERIENCE

Jan. 2020 – Current   Los Angeles, CA
Insight Data Engineering│ Insight
Project URL: https://github.com/tracy15932/tone-gauge

- Implemented a data pipeline by ingesting over 1.5TB/yr data with Spark and MySQL to create dashboard using Tableau to analyze news tone trend over period of time.
- Helped investors and marketing teams to understand the global tone trend on news event in order to make more profitable decision.

Dec. 2018 – Dec.2019   Pasadena, CA
Junior Database Administrator | OD International Investments Inc.

- Performed troubleshooting, administration, and configuration on MySQL databases for a third party B2B online payment platform and a social educational platform.
- Collaborated with team members to create database from QA to production stages and ensured robustness.
- Managed user access priorities and security issues.

Apr. 2018 – Aug. 2018   Taipei, Taiwan
Website Project Manager Intern | DD Studio

- Managed project schedules, tracked deliverables, and established priorities for the team.
- Collaborated and negotiated with third-party software partners and internal teams to provide reliable solutions to client's problems.
- Created Gantt Diagram, Site-Map, and tested the web to resolve any rising problem on the function of the website.
- Evaluated the trade-off for web development methodologies to improve overall performance by 30%.

PROJECTS

Salary Prediction Project
Project URL: https://github.com/tracy15932/salary_prediction

- Predicted salary based on job features which can help HR department to improve recruitment process.
- Used Python to manipulate about 1 million entries of data to perform data cleaning, visualization and analysis.
- Developed machine learning models using Multiple Linear Regression, Random Forest, and Gradient Boosting to forecast the best compensation strategy by minimizing the Mean Squared Error (MSE).
- Achieved an MSE from 470 at the baseline model to 357 under cross validation.

Monopoly Game Board Simulation Project

- Used R to implement 1000 simulation of a two-player game and visualized the landing frequencies for each space with histogram to analyze the distribution.