Runda Tian

1913 Berkeley Way, Berkeley, California 94704
510-384-0770 tian1998@berkeley.edu linkedin.com/in/rundatian github.com/8and24and10

Education

University of California, Berkeley **Aug. 2021 – Dec. 2022**
Master of Engineering, System Engineering, GPA: ***3.68*** *Berkeley, California*
Relevant Coursework: Data Structure and Algorithm, Full Stack Development, Principles and Techniques of Data Science
Certificates: Graduate Certificate of Applied Data Science, Certificate in Entrepreneurship & Technology

Chengdu University of Technology **Sep. 2016 – Jun. 2020**
Bachelor of Engineering in Civil Engineering, GPA: ***3.81****, ranked* ***2/223*** *of the class* *Chengdu, China*
Honors and Awards: National Scholarship, Outstanding Graduate Awards, 2nd Place in Mathematical Contest in Modeling

Experience

Precise Mind, Inc **Sep. 2021 – Dec. 2021**
Software Engineer Intern *Berkeley, California*
- Analyzed historical financial data from the 1990s, including financial statements, earning calls, fundamentals of companies and stock price momentum structure for institutional investors and retail investors on different marketing time frames.
- Extracted real-time stocks information from Twitter API to build pipeline in MangoDB.
- Applied sentiment analysis on online news for the performance of ARIMA analysis and implemented LSTM on the historical stock price to clean the data and predict the movement of stock price.
- Constructed dashboard and documented business reports on the investment plan based on the data visualization.

Deloitte Consulting **July. 2020 – Oct. 2020**
Data Analyst Intern *Chengdu, China*
- Participated in the data standard implementation, management and data system planning of Sichuan Bank.
- Designed workflows in Alteryx to automate work processes including data transformation and reports generating.
- Improved the performance of SQL queries and mapping logic which cut the data reports running time by 20%.
- Assisted in creation of weekly data reconciliation templates by Python for a financial product category, and visualized important business metrics in dashboards using Tableau.

Projects

COVID-19 Data Analysis Through Modelling | *Scikit-Learn, Seaborn, NumPy* **Nov. 2021 - Dec.2021**
- Applied data cleaning, exploratory data analysis, and time series analysis techniques in Python on public COVID-19 datasets released by John Hopkins Center for Systems Science and Engineering (CSSE) and the CDC.
- Developed a feature engineering pipeline by using Lasso cross-validation to select best COVID-19 features to be used in the machine learning modeling and prediction process.
- Implemented machine learning algorithms including logistic regression, decision trees, and random forest to predict county-wide CDC social mobility index percentile, achieving of 95% with selected COVID-19 features.

Scheme Interpreter | *Scheme, Python, SQL* **Nov. 2021**
- Implemented an interpreter for a functional subset of the Scheme programming language.
- Designed and achieved interpreter's function that can tokenize, evaluate, and compile various inputs in Python.
- Programmed a recursive Art Contest using advanced python data structures such as trees, hash, and graphs.

Spam/Ham Classifier | *Numpy, Pandas, Matplotlib* **Nov. 2021**
- Detected ham/spam messages from natural language data which contains 5574 short message service.
- Applied bag-of-words approach to transform the original corpus to a vector and created a pipeline as an estimator API.
- Trained and tested the classifier using the Naive Bayes classifier algorithm and achieved 97% detection accuracy.

Enigma | *Java* **Feb. 2022**
- Programmed a Java-based advanced simulation of the Enigma encryption machine used in World War II.
- Accommodated different machine settings including various alphabet permutations, number and setting of rotors, and plugboard pairings to encrypt and decrypt given messages.

Technical Skills

Languages: Python, Java, C, Scheme, Git, HTML/CSS, JavaScript, SQL
Developer Tools: VS Code, Eclipse, Google Cloud Platform, MangoDB
Technologies/Frameworks: Machine Learning with Pandas, NumPy and Scikit-learn, GitHub, Computer Aided Design