

Sun, Liyuan

San José, CA 95133 | (669) 265-8218 | sunliyuan8@gmail.com |
[linkedin.com/in/liyuan-sun/](https://www.linkedin.com/in/liyuan-sun/) | github.com/liyuan789 | leetcode.com/liyuansun/

SUMMARY

Seeking a Software Engineer or Data Scientist position. M.S. Statistics new graduate with solid coding and analytical skills in web visualization, data structure, algorithms, and data mining. Expert in Python, Java, web development tools, database (SQL, NoSQL), and machine learning, offering hands-on debugging, trouble shooting, problem-solving and leadership, business skills.

WORK EXPERIENCE

Quantitative Researcher, Ethereum Foundation, CA Aug. 2019 – Jan. 2020

- Outlined the Ethereum 2.0 solution to the data availability problem. Performed game-theoretical analyses of the early key reveal mechanics to incentivize the participants to act honestly and make Ethereum Foundation site more secure.

Data Science Intern, Inspur System, China Apr. 2018 – Sep. 2018

- Built data pipelines to integrate data from 3rd-party providers, consolidated 200k+ entries clinical data spanning 5 years
- Developed machine learning algorithms to simulate the morbidity of stroke and cardiovascular diseases improving R2 score by 15.65% compared to rule-based model. (**Python, SQL, Tableau, VBA**)
- Created customizable **Tableau** dashboard with multiple filters, presented to 100 key stakeholders, boosted usability by 50%

PROJECT EXPERIENCE

Event Search and Ticket Recommendation / Java web service development Jan 2020 – Mar 2020

- Developed an interactive webpage for events search with **AJAX** technology (**HTML, CSS, and JavaScript**) and implemented recommendation based on search history and favorite records.
- Created Java servlets (Tomcat Server) with **RESTful** APIs to handle HTTP requests and responses
- Built relational databases **MySQL** use **MAMP** to capture real business data from Ticketmaster API.
- Deployed server-side to **Amazon Web Services** (AWS) EC2 to handle 150 queries per second tested by Apache JMeter.

Customer Churn Prediction in Telecommunication Industry / Machine learning Sep 2019 – Dec 2019

- Developed algorithms for telecommunication service vendors to predict customer churn probability based on labeled data (7k entries) via Python programming
- Trained supervised machine learning models including **logistic regression, random forest, KNN, SVM, Neural Network, and Ensemble Model**; tuned hyper-parameters with **RandomizedSearchCV**
- Evaluated model performance via key metrics of the confusion matrix and obtained best **ROC-AUC** score at 0.845.

Stock Data Exploration Website Design / Python web page design Dec 2018 – Dec 2018

- Led a team of three to build an interactive web page for user to compare stock performances with earnings and interests using **Python, Highcharts.js, D3.js, and JavaScript**
- Scraped stock data from IEX and **Quandl APIs** to handle HTTP requests and responses
- Stored target data to **SQLite** and loaded to the **Flask** route to process the visualization based on user input

EDUCATION

San José State University, MS in Statistics Aug 2017 – May 2020

Achievement: Top 10% students selected in CAMCOS project;

Scholarship: 2020 WIA (Women in Analytics) Student Scholarship

Shandong University of Finance and Economics, BS in Accounting Sep 2011 – Jun 2015

Achievement: Top 3% students; 2015 Outstanding Graduates Awards of Shandong Province

SKILLS

Programming Languages: Python (NumPy, Pandas, Scikit-Learn), Java, R, SQL/NoSQL, HTML/CSS, JavaScript (D3.js, Highchart.js, Bootstrap, AJAX)

Tools: Eclipse, Jupyter Notebook, MATLAB, JMP, Tableau, MS Excel, Flask, AES, Github, Hadoop, Spark