AOYUAN (EVE) LIAO

Irvine, CA

917-530-9569 ⋄ eveliaocc@gmail.com ⋄ www.linkedin.com/in/eveliao ⋄ github.com/liaoaoyuan97

PROFESSIONAL EXPERIENCE

Othering & Belonging Institute at University of California, Berkeley Oct 2021 - Present
Data Scientist *Remote in Irvine, CA*

Othering & Belonging Institute is a policy research institute that aims to address social exclusion and inequality in public policymaking. I work with a small team of data analysts to conduct quantitative analysis and provide evidence-based insights.

- Built a large-scale data processing pipeline to extract and clean over 200 million property tax records across California. Examined the racial disparity impact of Proposition 13 by applying regression, correlation analysis, and descriptive and inferential statistics.
- Created a reusable Python module to extract and tailor indicators from Census Bureau data as per demand. This module eliminated redundancy by 80% in the implementation of the community condition study.
- Analyzed national home mortgage data and created data visualizations to investigate racial discrimination in mortgage lending.
- Designed and updated Inclusiveness Index, one of the first metrics in academia, which holistically examines the degree of inclusivity experienced by marginalized groups across the globe and within the US.
- *Languages:* **Python, R** | *Analytic Tools:* **StatsModels, Pandas, Excel** | *Data Visualization Tools:* **Matplotlib, Seaborn, Plotly** | *Mapping:* **GeoPandas, Leaflet**

Open Source Contributor Sep 2020 - Present

- ***Pandas*** – Refactored code to enhance its readability and accelerate the performance of user-facing APIs.
- ***Spark*** – Updated Spark SQL documentation to promote new features and inform users of changes in the latest release.

Loadsmart Inc. Mar 2017 - Nov 2019
Data Scientist *New York, NY*

Loadsmart is a digital freight marketplace and fleet management platform, which automates how freight is priced, booked, and shipped. I was primarily responsible for providing machine learning solutions to its two core features: freight pricing and carrier matching. I contributed to their backend infrastructure, ML models, and data processing pipelines.

- Built, evaluated and optimized the predictive pricing model for full truckload shipping to forecast price fluctuations due to seasonality, supply vs demand, and geographic information via ML algorithms.
- Developed Pricing REST API in Flask, which handles over 270,000 quotes per day across the US. Migrated its feature transformation code base from Python to Golang, reducing the average response time by over 50%.
- Launched the suggested carrier ranking model machine-learned from historical carrier responses to routes, increasing the mean average precision at top 25 by over 10%.
- Designed and implemented A/B experiments that increased quote conversion rates by 20% and enhanced truck availabity email response rates by 15%, in collaboration with product, operations, and marketing teams.
- Architected KPI dashboards and automated reports to provide executives with real-time business performance insights.
- *Languages:* **Python, Golang, SQL** | *Database:* **Redshift, Postgres** | *Analytic Tools:* **Pandas, Spark** | *Data Visualization Tools:* **Tableau, PowerBI** | *Modeling Tools:* **Sklearn, TensorFlow, PyTorch** | *ETL Tools:* **AWS Data Pipeline, Apache Airflow** | *Deployment Tools:* **Docker, Kubernetes** | *Web Framework:* **Django** | *Computing Cloud:* **AWS**

Snapwiz Inc. Jun 2016 - Aug 2016
Data Science Intern *Fremont, CA*

Snapwiz is the owner of GLIDER.ai, an AI-powered recruitment platform for screening and interviewing candidates online. I worked closely with a cross-functional team on its candidate prescreening and recommendation system.

- Leveraged Natural Language Processing algorithms to parse texts in candidates' resumes both syntactically and semantically, recognize entities such as job titles, and detect proportions of multiple topics from work experience descriptions.
- Created interactive analytic dashboards for inspecting unfair candidate match scores, and fixed the corresponding bugs.
- *Languages:* **R, Python, Unix Shell, HTML** | *Database:* **MongoDB** | *Modeling Tools:* **NLTK, SpaCy, Word2vec** | *Web Scraping Libraries:* **Beautiful Soup, Scrapy** | *Computing Cloud:* **Microsoft Azure**

EDUCATION

Columbia University in the City of New York *Sep 2015 - Dec 2016*
M.A. in Statistics — Overall GPA: 3.8 *New York, NY*
Beijing Institute of Technology *Sep 2011 - Jun 2015*
B.Eng. in Automation (Electrical Engineering) — Overall GPA: 86.22/100 *Beijing, China*
B.A. in Economics — Overall GPA: 83.35/100 *Sep 2013 - Jun 2015*