

Xuezhan (Nik) Sun

1210 S Indiana Ave, Chicago, IL 60605 | 805-403-9974 | nikswenzx@gmail.com | <https://kugelbrecher.github.io/portfolio/>

EDUCATION

Rice University, Houston, TX

Dec. 2023

M.S. Data Science, GPA: 3.92/4.00

University of California Santa Barbara

Jun. 2021

B.S. Mathematical Sciences, B.A. Economics, Minor German

Graduation with Honors | Distinction in Economics Major | GPA: 3.71/4.00

SKILLS

Languages & Frameworks: Python, SQL, TypeScript, React, Flask, Node.js, PyTorch, TensorFlow, HTML/CSS, Java, Spark

Tools & Platforms: AWS, Git, SAS, Google Cloud, Tableau, Excel, STATA

Core Competencies: Web Development, Data Analysis, Machine Learning, NLP, Computer Vision, Data Mining, Visualization

WORK EXPERIENCE

Full-Stack Data Scientist | Prospect Equities, Chicago, IL

Feb. 2024 – Present

Dashboard Website for Real Estate Property Evaluation (**Python, AWS, PostgreSQL, React, Flask, GCP**)

- Architect and build database to efficiently manage large-scale real estate data from Multiple Listing Service API, with **real-time** replication. Augment data with scraped rental listings from various sources to include non-agent listings.
- Conduct in-depth modeling on the historical data to forecast house prices referring Zillow public patent methodology and predict **likelihood** to sell for all Midwest properties, improving investor feedback and engagement.
- Streamline report accessibility through address search and **draw map search** for key metrics like capitalization rate and annualized return with sorting, filtering, visualizing, and downloading features, cutting **80%** communication overhead.
- Enhance website security against scrapping and potential attacks by parameterizing SQL queries, WAF, encrypting database and obscuring API and sensitive code from public exposure.

Performance Engineer Intern | Silicon Ranch Corporation, Nashville, TN

May. 2023 – Aug. 2023

- Developed a database using **Python**, featuring automatic daily and weekly updates; designed intuitive user interfaces for customized data extraction and analysis, resulting in a **90%** decrease in data preparation and processing time.
- Provided presentations and reports to communicate analysis findings with internal board utilizing **Tableau**, simplifying complex data insights into easily understandable and actionable information; delivered over 5 actionable actions.

Machine Learning Engineer Intern | Consentify, Seattle, WA

May. 2023 – Aug. 2023

- Collaborated in the design and development of a B2B data privacy and consent management system using **Typescript** and **React**, which features straightforward manipulation of customer consent choices.
- Fine-tuned and deployed a **Chat Bot** leveraging **gpt-35-turbo**, enabling clients to make API requests via chat, streamlining user experience and interaction.

Research Fellow | Rice, Houston, TX

Sept. 2022 – May. 2023

- Mined and integrated data from five distinct government open-source databases, enriching data repository and achieving a **40%** improvement in efficiency for future data cleaning processes through automation.
- Investigated the impact of U.S. fiscal policy during recessions, developed labor force indices with a **97%** accuracy rate compared to official data, and constructed a two-stage least square regression model to mitigate bias.
- Delivered **presentations** at Data2Knowledge Lab and Federal Reserve.

Operational Analyst | Fuhe New Materials, Zhenjiang, China

Oct. 2021 – Jul. 2022

- Secured 3 construction projects from **nationwide** corporations through accurate quote submission and bid strategies. Negotiated with suppliers and banks for cost-effective purchases and enhanced lending options.
- Managed daily raw material inventory using Excel and Kingdee, ensuring uninterrupted production despite unexpected transportation bans under government regulations.

RELEVANT PROJECTS

Capstone: Diagnose Diastolic Dysfunction in Pediatric Patients | Texas Children Hospital (TCH)

Aug. 2023 – Dec. 2023

- Image classification: image quality, imbalance data, tree models, Convolutional Neural Network, Grad-Cam interpretation ([link](#)).

Landmark Detection for Cephalometric Images | Rice, Houston, TX

Jan. 2023 – May. 2023

- Point recognition on CT images: U-Net model, 74% accuracy, visualization.

Blog Web App | Rice, Houston, TX

Aug. 2022 – Dec. 2022

- Full-stack blog web application: MongoDB, avatar store in Claudinary, RESTful API, OAuth, unit-testing.

Senior Thesis: Causality between Covid-19 and Smoking Reduction | UCSB, CA

Sept. 2020 – Mar. 2021

- Difference-in-difference casual effect analysis ([link](#)).