Taolan Jiaohaer, M.S.

(858) 568-1191 | jtaolan@ucsd.edu | Personal Website

EDUCATION

University of California, San Diego | San Diego, CA *Aug.2022-Jun.2024*
M.S. in Computational Social Science

University of California, Berkeley | Berkeley, CA *Jan.2021-May.2021*
Exchange Undergraduate Student

Nanjing University | Nanjing, China *Sep.2018-Jun.2022*
Bachelor of Law. in Political Science

SKILLS

Data Science Libraries: NumPy, Pandas, Scikit-Learn, TensorFlow, Pytorch, KenLM, OpenFST,
Statistical/Machine Learning Models: Generalized Linear Model, Time Series Analysis, Logistic Regression, Decision Tree, Random Forest, K-means, Gradient Boosting, Support Vector Machine, Neural Network,
Tools: Python, SQL, HTML, Julia, R, Tableau, Stata, Git, QGIS, ArcGIS
Certifications: IBM Databases and SQL for Data Science with Python. CITI Good Clinical Practice
CITI Certified Biomedical Research-Basic/Refresher

EXPERIENCE

Database Administrator and Visualization Specialist | MACRO Lab UCSD *Sep. 2023-Present*
• Creating, designing, and developing the MySQL database server for UCSD Macro Lab; using PyTorch to perform statistical analysis of the media data in the database.
• Building, monitoring, and maintaining Tableau dashboards. Making data analysis and visualization dashboards using lab data to conduct research on consolidation and financialization research in the media industry.

Computational Linguist Internship | Sensory Inc *Oct.2022-May.2023*
• Using open-source Finite-State Transducers (FST) toolkits (OpenFST and KenLM) to convert phonological rules as regular expression pattern substitution rules, into Weighted Finite-State Transducers (wFST) with Python。
• Building the FST from phonological rules, inverting and composing it with the other portions of the full architecture.

Machine Learning & Social Science Projects

Media Consolidation and Democracy | MACRO Lab UCSD *Spring 2024*
• Collecting, pre-processing, and cleaning the data from LSEG workspace as well as lab confidential data; converting data from Google Sheets format to MySQL databases; Conducting statistical analysis with the data including correlation, regression with machine learning models such as linear regression, non-linear regression with neural network. • Project page

Diabetes Prediction Using Machine Learning Algorithms | UCSD *Winter 2023*
• Train multiple machine learning algorithms such as logistic regression, SVM, and decision trees to perform diabetes prediction using handcrafted features based on people's characteristics (e.g. age, gender, etc.) using PyTorch and Skikit-Learn. • Project page

Wave Height Prediction with Dimensionality Reduction | UCSD *Fall 2022*
•Using PyTorch and publicly available data from NOAA to predict potential surf zone wave height in Hawaii North Shoreline with an Ensemble of Neural Networks and Regression algorithms. Project page.

The Ancient Silk Road Research | University of Chicago *Fall 2020*
•Conducted in-depth research on early Indian Ocean trade, focusing on The Periplus(1st-centure CE Roman merchant's report). Analyzed authorial intent, arguing the text served as a merchant handbook rather than a simple trade record. Examined transaction costs and regional commodity exchanges to understand economic dynamic. Wrote a paper demonstrating critical historical analysis and textual interpretation.

United Nation Talent Training Program | Nanjing University *Fall 2019*
•Completed all course of Nanjing university International Organization Talent Training Program.