

Zhexu (Kelvin) Li

Phone: 805-618-5607 • Email: zh1411@uscd.edu

I have experienced firsthand the explosive growth of data science applications in almost every aspect of my daily life. What I have learned as a Data Science major at UCSD truly opened my eyes to the tremendous potential of Data Science. It also made me eager to pursue graduate studies in order to further my knowledge and skills.

EDUCATION

University of California, San Diego

September 2017-June 2021

BS in Data Science. Cumulative GPA: 3.57

Core Courses: *Systems for Scalable Analysis, Machine Learning, Theoretical Foundations of Data Science, Practice of Data Science, AI: Probabilistic Models, Data Management, Data Analysis and Inference, Data Mining, Data Visualization, Spatial Data Science and Application.*

Honors: Provost Honors in Fall 19, Winter 20, Spring 20, Fall 20 and Winter 21

University of California, San Diego

September 2021-June 2023

MS in Machine Learning and Data Science (graduation expected in June 2023).

PROGRAMMING SKILLS

Proficient in Python (Pandas, GeoPandas, Scikit-learn, Dask, PySpark, ArcGIS), SQL(PostgreSQL), JavaScript

Familiar with R, Java, MATLAB, STATA

PROFESSIONAL PROJECTS

Indoor Airborne Infection Risk Estimator

Fall 2020-Winter 2021

Mentor: Professor Rajesh Gupta, Founding Director of Data Science Institute and Former Chair of CSE Department in UC San Diego.

We developed an indoor airborne COVID – 19 infection risk estimator. Our estimator features great transparency and extensibility while maintaining the ease of use. Algorithm was built based on solid research, normal users can easily use default settings to get decent estimations, while advanced users can import their custom rooms and modify assumptions based on the situation in their own regions. We also provide an interactive visualization for convenient comparison between cases. I developed the entire frontend of the website, and the majority of its backend which is a dash-based application. Visit the application website at: <https://hinanawits.github.io/DSC-180B-Presentation-Website/>.

ACADEMIC PROJECTS

Project on System for Scalable Analysis

Fall 2020

Analyzed large Amazon products and reviews datasets on cloud computing platforms including AWS using large-scale data processing libraries Dask and PySpark; Optimized codes to minimize time and space costs.

Project on AI: Probabilistic Models

Summer I 2020

Implemented several Python applications from scratch using Bayesian network, Hidden Markov Network and Markov Decision Process.

Project on Data Mining

Spring 2020

Analyzed Airbnb datasets and trained several machine learning models to predict rent price using Multi-layer Perceptron and Random Forest Regressor.

Project on Spatial Data Analysis

Fall 2020

Analyzed potential factors for bikes related accidents in San Diego County using GeoPandas.

Analyzed Wildfire Hazard Potential in San Diego County using ArcGIS Python API and performed classification of damaged buildings in imagery using Deep Learning Feature Classifier in ArcGIS Pro.

Predicted U.S presidential election result using geospatial and census information of U.S counties.