Jasper M. Dong

(925) 389-2017 | jasper.dong.2003@gmail.com | www.linkedin.com/in/jmudong

Education

University of California - Davis *Expected Graduation: June 2025*
B.S. in Statistics - Statistical Data Science Track *GPA*: 3.813
Minor in Geographic Information Systems
Honors: Dean's List (Fall 2021, Spring 2022, Fall 2022)
Activities: Davis Data Science Club, HackDavis
Relevant Coursework: Applied Statistical Methods, Statistical Data Science, Data & Web Technologies, Statistical Learning, Multivariate Data Analysis, Mathematical Statistics, Discrete Mathematics, Applied Linear Algebra, Spatial Data Analysis, Geographic Information Systems, Remote Sensing

Experience

Research Assistant - *Computational Climate and Ocean Group* *Summer 2024 - Present*
Advisor: Maike Sonnewald
- Performed exploratory data analysis and data preprocessing on large data sets (~1.2 million observations) using Python libraries and high performance computing tools
- Combined the use of manifold learning and agglomerative clustering to identify unique ocean regions of interest based on different physical parameters
- Implemented a point-based Chi-Square Test of Independence to ensemble clustering results for use as a clustering parameter validation method

General Member - *Davis Data Science Club* *2022 - 2024*
- Worked with peers on quarter-long projects involving data analysis and implementation of machine learning methods to summarize
- Gained exposure to the application of basic data science tools through participation in club workshops

Projects

HouseCat - *HackDavis* *2024*
- Worked with a team to implement web-based database application displaying local unofficial housing listings
- Worked on setting up Node.js and React.js interfaces to connect with MongoDB database containing web-scraped housing listings from community Facebook housing groups
- Presented and demonstrated prototype to judges of fellow students and industry professionals

Climate Change Analysis - *Davis Data Science Club* *2023*
- Worked with a team to explore dataset of global temperature patterns using pandas and numpy libraries
- Used matplotlib and seaborn libraries to create correlation plots and heatmaps for EDA visualizations

Skills

Programming Languages: Python, R, MATLAB, SQL, regex, XPath, HTML, CSS, JavaScript, LaTeX
Skills: EDA, Data Analysis, Machine Learning, Model Validation, Hypothesis Testing
Software: SLURM, Tableau, MongoDB, ArcGIS Pro, Google Earth Engine
Spoken Languages: English, Chinese, Japanese, Dutch