Jialin Wang

Email: jialin.wang@berkeley.edu
LinkedIn: linkedin.com/in/jialin-wang-a4a0b919b/
Mobile: +86-15861196829

EDUCATION

University of California, Berkeley — Berkeley, USA
Master of Analytics, Analytics — *August 2023 - July 2024*

New York University Shanghai — Shanghai, China
Bachelor of Science, Mathematics and Data Science (Mathematics Concentration) — *September 2019 - May 2023*
GPA: *Overall GPA: 3.83, Primary Major GPA: 3.91*
Honors: Major Honors in Mathematics (top one in Mathematics major), *Cum Laude, University Honors Scholar, Dean's List for Academic Year (Fall 2020 - Spring 2021, Fall 2021 - Spring 2022, Fall 2022 - Spring 2023)*
Courses: *Linear Algebra, Probability and Statistics, Mathematical Statistics, Multivariable Calculus, Mathematics Foundations of Machine Learning and Data Science, Machine Learning, Data Structures, Databases, Optimization, Ordinary and Partial Differential Equations, Discrete Mathematics, Numerical Analysis, Math Modeling, Econometrics*
Computer Science: *Python, R, SQL, MATLAB, HTML, MkDocs, Stata, Git, Wind, Excel, Word, PowerPoint, LATEX, Adobe Photoshop, iMovie*

RESEARCH AND PROJECTS

Analysis for Fourier Neural Operator and its Applications — Independent Researcher
Mathematics Senior Thesis — *January 2023 - May 2023*

- **Content**: In complex PDE settings, the neural operator, an efficient data-driven approach that aims to find the solution operator of PDEs, might be used. In contrast with traditional neural networks, which learn function mapping between finite-dimensional spaces, neural operators expand this learning to include operators between infinite-dimensional spaces. This allows for zero-shot generalization to higher-resolution evaluations and frees the neural operator from the grid's resolution and size for training data. Furthermore, if we rely on Fourier spaces for our training procedures, our solution training process will be more efficient.
- **Result**: In comparison to conventional numerical approaches, the Fourier neural operator has quasi-linear time complexity, allowing it to solve PDEs much more quickly. Numerical experiments are performed on Darcy Flow as well as Fokker-Planck equations to prove such properties of the Fourier neural operator.
- **Instructor**: The instructor is Professor Mathieu Laurière from NYU Shanghai.

A Multi-stakeholder Book Recommendation System for Libraries — Independent Researcher
Data Science Senior Capstone Project — *September 2022 - December 2022*

- **Content**: Processed twenty thousand book abstracts using the TF-IDF method and GloVe and compared their performances, to construct book and student profile vectors. Used LASSO to predict the ratings and popularity of a set of new books. Turned the classical consumer-centric model into a multi-stakeholder, multi-objective combinatorial optimization problem and implemented Genetic Algorithm to find desired books for students, and promote the lending of under-exposed books simultaneously.
- **Result**: Developed a book recommendation system that offers students desired books while expanding the library's loan volume of various books, achieving a near-even distribution of books lent to maximize the allocation of public recourse. The invented system reduced the variance of books' lending times by 16 percent compared with the single-objective recommendation.
- **Instructor**: The instructor is Professor Hongyi Wen from NYU Shanghai.

The Application of Matrix Completion and PCA in the Movie Recommendation System — Independent Researcher
2022 New York University Shanghai "Deans' of Undergraduate Research Fund" — *June 2022 - September 2022*

- **Content**: Conducted a theoretical analysis and research of matrix completion, Robust PCA, Sparse PCA, and related solving algorithms including Singular Value Thresholding, Augmented Lagrange Multiplier, and Dual Method, to recommend movies for customers using Matrix Completion and Collaborative Filtering. Performed real data implementation using MATLAB to test the different models.
- **Result**: Successfully conducted the experiment using data collected from my friends and the official Netflix movie dataset after data cleaning, recommended movies to them, and then compared the results given by three algorithms. Finally, ALM had the smallest nuclear norm and was the fastest algorithm.
- **Instructor**: The instructor was Professor Shuyang Ling from NYU Shanghai.

Predict National Life Expectancy using Lasso and Comparison between other Models — Team Leader
Project of course "Machine Learning" — *April 2022 - May 2022*

- **Content**: Performed data cleaning and normalization for over 4000 rows of data using Python, filling missing values of features (with less than 30% missing) using the average. Predicted life expectancy in different continents for the past two decades through the analysis and comparison between different regression models such as LASSO, Ridge Regression, Regression Tree, and Random Forest Regressor. Tested the models on the WHO national life expectancy dataset and did cross-validation.
- **Result**: Applied the model with the best prediction performance (R-square 99.5%) and interpretability: LASSO, to do longevity prediction and explained the selected important features for each continent. We got full marks for the project.
- **Instructor**: The instructor was Professor Mathieu Laurière from NYU Shanghai.

Post-selection Inference for Lasso   Team Member
Project of course "Mathematical Foundations for Data Science and Machine Learning"   *March 2022 - May 2022*

- **Content**: Studied how to form valid confidence intervals for the coefficients of the variables selected by Lasso. Performed a theoretical analysis and broad literature reviews on how to characterize the selection event obtained by Lasso, and then computed the post-selection confidence interval. We ran a diabetes dataset experiment to visualize the post-selection confidence interval. Moreover, to better understand the validity of the obtained post-selection confidence interval, we ran simulations under low-dimensional data and high-dimensional data.
- **Result**: Concluded the post-selection methods for theoretical and practical use and proved its advantages compared to classical approaches such as Bayesian methods and the bootstrap. The confidence interval generated by post-selection methods is more likely to cover the ground truth.
- **Instructor**: The instructor was Professor Shuyang Ling from NYU Shanghai.

Finding the Best Parameters for Heston Model under Chinese Stock Market   Team Leader
2021 New York University Shanghai "Deans' of Undergraduate Research Fund"   *June 2021 - November 2021*

- **Content**: Trained the best parameters for the Heston model using Chinese stock market ETF-50 data in 2021 through the Bisection method and the Simulated Annealing Algorithm, did back-testing, and finally gave an error analysis. As the team leader, provided technical support for teammate, kept in touch with the mentor, and scheduled meetings to discuss the project's progress.
- **Result**: The predicted value was close to the actual stock value, with an error of fewer than 0.13 units. Our team won "The Best Research Project of STEM and Media" in the Fall 2021 NYU Shanghai Undergraduate Research Symposium.
- **Instructor**: The instructor was Professor Wei Wu from NYU Shanghai.

INTERNSHIP EXPERIENCE

NYU Shanghai Mean Field Games Research Group   Shanghai, China
Research Assistant   *June 2023 - Now*

- **Code Cleaning and Implementation**: Under the supervision of Professor Mathieu Laurière at New York University Shanghai, cleaned, checked and updated codes for numerical methods of mean field games and mean field control using Python for the tutorial purpose. Generated experimental results and reproduced figures in the scientific papers, and compared the results with other deep learning approaches. Wrote Newton method Python codes for different dimensions and boundary conditions of PDEs, and reproduced the congestion experiment.

Discover Financial Services   Shanghai, China
Data Analyst Intern   *June 2023 - Now*

- **Front-end Development and Data Analysis**: Developed tutorial websites on Data Analytics for new-hired training using Github pages and MkDocs. Worked with teammates within the Credit Card Management Assistance group to develop predictive models, conduct simulations, perform ad-hoc analysis, design and produce analytic reports to solve business problems. Communicated with colleagues in the US to deal with technical problems.

Western Securities   Shanghai, China
Data Engineer and Analyst Intern   *February 2023 - May 2023*

- **Data Engineering**: Under the SQLAlchemy structure, stored the data from external suppliers in a normalized format using Python and MySQL. Analyzed raw data and features, designed database table structures according to the analysis, and then created database tables using MySQL. Performed data cleaning and normalization using Python and wrote configuration files to achieve data storage process.
- **Data Analysis**: Applied various machine learning models such as time series analysis (LSTM), SVM, Random Forest, and Kmeans to study the financial data, interpreting the specific industry from the data provided. Developed and designed relevant API to realize convenient and fast financial data analysis and visualization services.

Shanghai Goldlake Fund   Shanghai, China
Quantitative Analyst Intern   *July 2022 - August 2022*

- **Quantitative Transaction Modeling Work**: Assisted in the construction of quantitative transaction strategies and related factor mining based on various industry materials and literature reviews worldwide, and summarized the strategies for the team.
- **Data Analysis**: Used Python to process past data and interpreted financial markets using historical data. Evaluated the stock and security data using risk indicators and other models such as time series analysis (ARMA). Constructed the alpha models and rate the effectiveness (Information Coefficient) of different alpha, presenting the performance of each alpha through graphs.