Chenghui Li

cli539@wisc.edu · (+1) 608-358-2030 · github.com/chl781

Education

University of Wisconsin-Madison, Wisconsin, U.S. Aug. 2018 – Present

M.S. in Data Science; Expected to May 2020; *GPA: 4.00/4.00*

Zhejiang University, Zhejiang, China Aug. 2015 – June 2019

B.S. in Mathematics and Applied Mathematics; *GPA: 3.77/4.00; Major GPA: 3.95/4.00*

Projects & Experience

Consistency of an Algorithm for Manifold Clustering Jan. 2019 – Present

Programming, Algorithm design and Theoretical proof Professor: Nicolas Garcia Trillos

- Proposed a methodology for manifold clustering to recover a low-dimensional representation of the data coming from several overlapping geometric structures.
- Try to develop rigorous mathematical theory establishing its consistency by using ideas from geometry, probability and calculus of variations.
- Used Matlab to verify the efficiency of the model on simulated data.

Yelp Business Analysis Oct. 2019 – Nov. 2019

STAT 628: Data science Practicum, Course project Professor: Hyunseung Kang

- preprocessed one hundred thousand business review data and used counts of words to fit a linear regression to find the relationship between the reviews and the scores.
- Used keywords to separate the texts into three parts, including service, price and flavor, and then used LSTM to predict the scores.
- Used Shiny with Html to present real-time response including the statistical tests, prediction of scores and advice to the business.

FFI Algorithm Performance in Bin-packing MinSum Problem Feb. 2018 – June 2019

Theoretical proof Professor: Zhiyi Tan

- Used the construction method to improve the performance ratio lower bound from 1.25 to 1.35.
- Improved the asymptotic performance ratio upper bound from 1.83 to 1.5 by using the ideas from the construction method and proof by contradiction.

High Dimensional Simulation in SPCA, PCA, LASSO and PLS Apr. 2019 – May 2019

Stat 471: Introduction to computational statistics, Course project Professor: Fangfang Wang

- Used the latent model to analyze SPCA's performance on prediction.
- Used R to generate multicorrelated data and compared SPCA algorithm's performance on data with others.

Modeling for Hematopoiesis Prediction Dec. 2018

STAT 601: Statistical Inference, Course project Professor: Zhengjun Zhang

- Compared the efficiency of GMC(generalized measure of correlation), a complement tool for correlation, with other classical algorithms including PCA, PLS, LASSO on the prediction of data.
- Interpreted the results and the parameters of these methods to compare their difference.

Artificial Intelligence and Industry Program in England Aug. 2018

- Visited Imperial College London, the University of Oxford and the University of Cambridge to meet with professors and listen to their lectures about applications in artificial intelligence.

Lagrangian Flux Calculation via Donating Volumes in 3 Dimension Oct. 2017 – Dec. 2017

Directed Study: Multigrid Method on Irregular Domains Professor: Qinghai Zhang

- Defined donating volume as a 3-dimensional generalization of donating region by using pathlines and streaklines, and then determined the generalized flux identity.
- Proved that donating volumes were index-by-index equivalent to the corresponding flux sets of the same volume by using ideas from geometry, fluid mechanics and homological algebra.

Volunteer Teaching Experiences

International Teaching Volunteer in Indonesia Feb. 2017

Teaching Volunteer in Yunnan province, China(twice) June 2016 & June 2017

Honors and Awards

- Visiting International Student Academic Excellence Award(twice) at Madison Mar. 2019 & July 2019
- Chu Kochen Honors Program at Zhejiang University June 2019
- Qiushi Pursuit Science Class(Major in Mathematics) at Zhejiang University June 2019
- Putnam Exam Top 200 2019
- First Class Scholarship for Elite Students in Basic Science at Zhejiang University Jan. 2017

Skills

- **Programming Languages**: Proficient in R, Python and Matlab, experienced in Latex, C and HTML
- **Languages**: English - Fluent, Mandarin - Native speaker

Miscellaneous

- **Personal website**: https://chl781.github.io/
- **Relevant courses**: Mathematical Statistics, Real Analysis, Topics in Probability, Multi-variate Statistical Analysis, Statistical Methods, Statistical Learning Theory, Optimal Transport Machine Learning, Stochastic Processes, Time Series, Regression Analysis, Point Set Topology, Differential Geometry, Combinatorial Optimization