LI LIU

THIS CV IS FOR RECORDING ONLY
Siming South Road 422, Xiamen, China
22920162204022@stu.xmu.edu.cn ⋄ liulidaily@163.com

EDUCATION

B.S. in Computational Mathematics (Honor Program) Spetember 2016 - June 2020
School of Mathematics, Xiamen University
Major GPA: 3.89/4.0; 91.67/100 Major Rank: 1/28
Math Courses: Numerical Analysis, Functional Analysis, Numerical Solutions of Partial Differential Equations, Numerical Linear Algebra, Probability, Mathematical Statistics, Regression Analysis, Operations Research and Optimization, etc.
Programming Courses: C Programming, Objected-Oriented Java, Algorithms Design and Data Structures
Pending Courses: Database Management, Numerical Optimization, Data Analysis and Matrix Computation

B.S. in Mathematical Economics September 2017 - June 2019
Dual Degree, WISE, Xiamen University
GPA: 3.78/4.0; 89.75/100 Rank: 1/11
Courses: Microeconomics, Macroeconomics, Econometrics, Time Series Analysis, Data Mining

PROJECTS

Modification in the accuracy of clustering for single-cell data in biostatistics June 2019 - August 2019
Supervised by Zhixiang Lin, Chinese University of Hong Kong

- Reproduced model-based Bayesian clustering model for single-cell data by considering the connection with gene and accessibility of different chromatic by MCMC with Beta prior in R.
- Exploited Old Faithful Data to test Dirichlet Process in Guassian and Uniform prior. Plotted simulated probability density function.
- Implemented Dirichlet Process with only accessibility chromatic data and simplified Bayesian statistical model to update cluster indicators by Split and Merge algorithm, achieved better accuracy than algorithm-based models such as K-means.

AI Workshop: Exploiting Feature Engineering and Artificial Intelligence to solve real problems Aug 2018 - September 2018
Organized by Senior Research Fellow Qing Zhu, UC Berkeley

- Grasped basic theroies and applications about statistical learning and deep learning, such as random forest and CNN, in a short time.
- Successfully distinguished cats and dogs by a CNN model with near 100k parameters in 10k images, achieved nearly 98% accuracy in the test data.
- Improved the ensemble learning structure to identify financial fraud based on Kaggle by more suitable features selection with the aid of advanced economic models.

Data Analysis Program: Constructing real-time CPI indices for Online E-commerical Platforms Nov 2016 - April 2017
Organized by Jiaming Mao, Xiamen University

- Collected the price indices of nearly 2M goods on different e-commercial platforms by Python.
- Organized data in MySQL and built query API for further economic research.
- Presented the visualization of the price indices by both seaborn and ggplot2.
- Taught group members how to use LaTeX, Git and basic Python in seminars.

Online Platform for Knowledge Sharing July 2017 - Present
Leader

- Created a personal blog on Zhihu (Chinese Quora) and uploaded nearly 100 articles about Mathematics, Statistics and Programming, serving as tutorial and resources for eternal learning. Attracted nearly 10000 followers and 900 thousand readings.
- Exploited Web Techniques to build a website and maintained as a platform for knowledge sharing in School of Mathematics, Xiamen University. Website :https://zhuanlan.zhihu.com/c_119426147
- Was invited to attend the National Conference held by Zhihu in Beijing, China as one special guest due to high quality of the articles.

RESEARCH

Design Statistical Learning Algorithms to Solve Classification Problems in Online Transfer Learning December 2018 - March 2019

Advised by Min Jiang, Xiamen University

- Assimilated theories of modern AI models such as GAN and Xgboost in the weekly seminar.
- Implemented a series of recent Online Learning and Transfer Learning algorithms based on SVM and passive aggressive(PA) algorithm.
- Created a novel model for solving classification problems in Online Transfer Learning based on Online Gradient Descent and Transfer Component Analysis in transfer learning.
- Achieved near 10% higher accuracy compared with existing algorithms by adapting the step-size of Online Learning with respect to the overall behavior of data, holding datasets unchanged.
- Paper writing in progress.

Optimize Dictionary Learning Problems by Quasi-Newton Methods in Manifold Optimization March 2019 - Present

Advised by Wen Huang, Xiamen University

- Scrapped image datasets by Python and reproduced results of Image Processing by Conjugate Gradient Method on the SPD matrix manifold.
- Made numerical stability analysis with Quasi-Newton Methods and tested the speed compared with other methods by MATLAB.
- Successfully transferred Quasi-Newton Methods on different image datasets such as the textile images, achieved approximately twice the speed of the original method.

ACTIVITIES

WISERclub: A Data Science Club September 2016 - Present

Chairman for Undergraduate Group. Organized by Haiqiang Chen and Wei Zhong

- Explored and shared knowledge about recent events about Artificial Intelligence in the weekly seminar.
- Served as the instructor in Data Mining course, teaching algorithms, applications and case studies in Business Analytics.
- Summarized Data Mining as a Final Project with Python and Web Techniques: https://weakcha.github.io/WISERCLUB-Final/

INTERNSHIP

Luckin Coffee: Feature Engineering in Recommendation System and User Persona Aug 2019 - Apr 2020 (Expected)

Yield Growth Department: Algorithm Engineer Assistant

- Reorganized order data for one quarter to perform Co-clustering with Augmented Matrices(CCAM) by Python.
- Provided user preference pattern as a feature of user persona in the algorithm system of the department for others reference.
- Exploited broadcast in numpy and numba.jit to accelerate (2000 times against as) the original code for the CCAM model.
- Generated order time data and tuned parameters of Xgboost with oversampling and weight for the data in user buying prediction.
- Achieved 84.2% accuracy and 97.6% precision, higher than before, for 17.3M users buying 1-100 times in the recent quarter.
- Uploaded model online for A/B Test, which showed 1.8% increase in daily sales and nearly $50000 increase in daily revenue.
- Exploited Naive Bayes and Non-parametric Dimension Reduction to do Gender Prediction for User Profile Construction.
- Use K-means and DBSCAN to create a clustering model for feature Price Sensitivity with recent orders and coupon records.
- (In progress) Assisting to reorganize deep learning Python package and fix bugs in the company based on TensorFlow 2.0.

AWARDS AND HONORS

- **Honor Program Scholarship**, Sept 2019
- **First Prize** in the 'Lingsheng Cup' Programming Context in Xiamen University, Nov 2016.
- **Silver Medal** in ACM-ICPC, Fujian Province. May 2019
- **Honored Prize (Admitted in Final, Top 12)** in S.-T. Yau College Student Mathematics Contest, Applied Math, May 2019
- **First Prize** in Chinese Undergraduate Mathematical Modeling Contest, September 2017.
- **Second Prize** in The Chinese Mathematics Competitions, Nov 2018.
- **Meritorious Prize (Top 6%)** in American Mathematical Contest in Modeling, January 2019.
- **315/500** in Computing Accreditation for Professionals (Top 4% in China), March 2017.
- **National Scholarship**(Top 1% Students in China), September 2017.
- **'The best 100 papers'** Award in the National 'Surveying China' Paper Contest (1073 altogether),

SKILLS

Programming Skills	Proficient: C++, MATLAB, Python, SQL Understanding: R, Spark, Hadoop, Java, Julia, Scala, Web, Linux
Language Skills	Mandarin, English (TOEFL: 105, GRE : V161+Q169+AW3.5), Cantonese
Hobbies	Body Fitting, Badminton