Li Liu

22920162204022@stu.xmu.edu.cn | **Personal:** liulidaily@163.com

EDUCATION

Xiamen University (XMU)

B.S. in Information and Computational Science (Honor Program); Major GPA: 91.67 / 100; Rank: 1/28 — *Sept 2016 - June 2020*

B.S. in Mathematical Economics (Dual Degree); GPA: 89.75 / 100; Rank: 1/11 — *Sept 2017 – June 2019*

Courses: Numerical Analysis (96), Functional Analysis (95), Operations Research and Optimization(97), Numerical Solutions of PDE (97), Probability (95), Statistics (94), Applied Linear Models (88), Time Series Analysis (95), Financial Mathematics (92), Data Mining (89)

Pending Courses: Algorithms and Data Structures, Mathematical Statistics (Grad), Data Analysis and Matrix Computation (Grad)

PROJECTS

Data Science in WISERClub: Data Processing and Instructor Work in Data Mining — *Oct 2016 - Present*

Chairman, supervised by Professor Wei Zhong and Haiqiang Chen, XMU — *Xiamen, China*

- ✧ Collected the price indices of nearly 2M goods on different e-commercial platforms by Python and organized data in MySQL.
- ✧ 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/

Modification in the accuracy of clustering for single-cell data in biostatistics — *June 2019 – Aug 2019*

Research Assistant. Advised by Professor Zhixiang Lin, CUHK — *Hong Kong, China*

- ✧ Reproduced model-based Bayesian clustering model for single-cell gene and chromatic accessibility data by MCMC in R.
- ✧ Exploited Old Faithful Data to test Dirichlet Process in Gaussian and Uniform prior and plotted probability density function.
- ✧ Implemented Dirichlet Process with only accessibility chromatic single-cell data and Bayesian model to update cluster indicators.
  - ■ Solved by Split-Merge algorithm, achieved better accuracy than algorithm-based clustering models such as K-means.

2019 MCM/ICM B: Managed packing strategies and drone routes for medical service in Puerto Rico — *Jan 2019 - Jan 2019*

Leader, awarded Meritorious Winner (Top 6% in the world), COMAP — *Xiamen, China*

- ✧ Derived linear programmer to provide proper medical packing schemes to maximize the usage of the volume of container by Lingo.
- ✧ Abstracted map into an acyclic weighted graph and solved the shortest path for drones to fly by dynamic programming algorithm.

AI Workshop: Exploiting Feature Engineering and Artificial Intelligence to solve real problems — *Aug 2018 - Nov 2018*

Organized by Senior Researcher Qing Zhu, UC Berkeley — *Xiamen, China; Online*

- ✧ Distinguished cats and dogs by a CNN model with near 100k parameters in 10k images, achieved 98% accuracy in the test data.
- ✧ Improved the existing model on Kaggle to identify financial fraud with the aid of random forest and advanced economic models.

RESEARCH

Designed Statistical Learning Methods to Solve Classification Problems in Online Transfer Learning — *Nov 2018 – Apr 2019*

Advised by Professor Min Jiang, XMU — *Xiamen, China*

- ✧ Implemented novel Online Learning algorithms based on Passive Aggressive and other Transfer Learning algorithms.
- ✧ Created a novel model for solving classification problems in Online Transfer Learning.
  - ■ Based on Online Gradient Descent and Transfer Component Analysis.
  - ■ Achieved near 10% higher accuracy than existing algorithms, holding datasets unchanged.
  - ■ Adapting the step-size of Online Learning with respect to statistical characteristics of data. Paper writing in progress.

Optimized Dictionary Learning Problems by CG and Quasi-Newton Methods in Manifold Optimization — *March 2019 - Present*

Advised by Professor Wen Huang, XMU — *Xiamen, China*

- ✧ Scrapped images by Python, reproduced Dictionary Learning process by Conjugate Gradient Method on the SPD matrix manifold.
- ✧ Conducted image classification problem by KNN and one-to-all SVM with 5-fold cross-validation.
- ✧ Made numerical stability analysis with Quasi-Newton Methods and tested the speed compared with other methods by MATLAB.
- ✧ (In progress) Transferring Quasi-Newton Methods on different image datasets such as textile images and YouTube facial pictures.

INTERNSHIP

Luckin Coffee: Feature Engineering in Recommendation Systems — *Sept 2019 - June 2020 (Expected)*

Yield Growth Department: Algorithm Engineer — *Xiamen, China*

- ✧ 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.
- ✧ (In progress) Assisting to reorganize deep learning Python package and fix bugs in the company based on TensorFlow 2.0.

AWARDS AND HONORS

- ✧ **Silver Medal (15/136)** in ACM-ICPC, Fujian Province — *May 2019*
- ✧ **China Computer Federation (CCF) Student Member** — *May 2019*
- ✧ **Admitted in Final (Top 12 in China)** in S.-T. Yau College Student Mathematics Contest — *May 2019*
- ✧ **Second Prize (Top 5%)** in The Chinese Mathematics Competitions, China Mathematical Society — *Nov 2018*
- ✧ **National Scholarship (Top 1% Students in China)** — *Sept 2017*

SKILLS

- ✧ **Programming (Proficient to Understanding):** C++, LaTeX, MATLAB, Python, R, Java, SQL, Spark, Scala, Linux, Julia, Web
- ✧ **Language:** Mandarin, English (TOEFL: 105, GRE: 330+3.5), Cantonese