

EDUCATION**Xiamen University (XMU)***B.S. in Information and Computational Science (Honor Program); Major GPA: 91.83 / 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), Stochastic Process (94), Time Series Analysis (95), Financial Mathematics (92), Data Mining (89), Algorithms and Data Structures (93), C++ Programming (90), Objective-oriented Java (85), Numerical Optimization (97)**Self-Learning Courses:** Convex Optimization (CMU 10-725), Intro to Machine Learning (CMU 10-601)**WORKING EXPERIENCE****Luckin Coffee: Feature Engineering in Recommendation Systems***Sept 2019 - June 2020**Data Scientist (Machine Learning)**Xiamen, China*

- ✧ Reorganized order data for one quarter to perform Co-clustering with Augmented Matrices(CCAM) by Python.
 - 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.
- ✧ Assisted to reorganize deep learning Python package and fix bugs in the company based on TensorFlow 2.0.

Cardinal Operations: Supply Chain Management for Orion*July 2020-Sept 2020**Data Scientist (Machine Learning)**Beijing, China*

- ✧ Connected Sales, Inventory and Order data into a large table by primary key goodsID.
- ✧ Tested different machine learning algorithms, such as LASSO, Random Forest, Xgboost and LightGBM, to do sales prediction.
- ✧ Created and Packaged a simulation system for supplies and demands by Simpy in Python.

Didi: Machine Learning System to solve Risk Control Problems*Sept 2020-Nov 2020**Data Engineer (AI System)**Beijing, China*

- ✧ Coded data preprocessing operators with Spark and Flink and collaborated with leaders to design UI and backend to simulate data analysis procedures.
- ✧ Collected various datasets to test the stability of the system.

Kuaishou: CTR prediction in Live-streaming Module*Dec 2020-Mar 2020**Data Scientist (Deep Learning)**Beijing, China*

- ✧ Tested three-layers DNN, Transformer, DIN, DIEN, DeepFM with Resnet, Rezero, and multi-task learning (MMOE and PLE) to optimize models in CTR prediction.
 - Achieved total 0.7% increase in model AUC and nearly 100% increase in the overall CTR for the live-streaming module.
- ✧ Created new user, item and combined features and uploaded onto the real-time system, with which achieved 0.2% increase in model AUC.

SELECTED PROJECTS**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**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.
- ✧ 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.

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 Transfer Learning algorithms based on Passive Aggressive and other Transfer Learning algorithms.
- ✧ Created a novel model for solving classification problems in Online Transfer Learning.
 - Achieved near 10% higher accuracy than existing algorithms, holding datasets unchanged.

Optimized Dictionary Learning Problems by CG and Quasi-Newton Methods in Manifold Optimization*March 2019 - Present**Advised by Professor Wen Huang, XMU**Xiamen, China*

- ✧ Reproduced Dictionary Learning process by Conjugate Gradient Method on the SPD matrix manifold to solve de-noising and classification problems.
- ✧ Made numerical stability analysis with Quasi-Newton Methods and tested the speed compared with other methods by MATLAB.
- ✧ Transferring Conjugate Gradient and BFGS methods on Manifolds on different image datasets such as textile images and YouTube facial pictures.

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*
- ✧ **National Scholarship (Top 1% Students in China)** *Sept 2017*

SKILLS

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