Yumeng Cao

Tel: (857)-275-8825 | **E-mail:** yumeng88@bu.edu

Web: https://yumengccc.github.io/ | **GitHub:** https://github.com/yumengccc

EDUCATION

Boston University, School of Public Health Boston, United States

M.S. in Biostatistics 09/2022-Now

- Related courses: Probability, Mathematical Statistics, Linear Models, Applied Genetic Analysis (GWAS using R and bash), Genomics Data Mining and Statistics (R), Intermediate Statistical Computing and Applied Regression Analysis (SAS), Applied Statistics in Clinical Trials (SAS), Statistical Methods in Epidemiology (SAS), et al.

Zhengzhou University, School of Mathematics and Statistics Zhengzhou, Henan, China

M.S. in Applied Mathematics 09/2018-07/2020

- Related courses: Basic Algebra, Basic Topology, Combinatorial Mathematics, Functional Analysis, Differential Manifold, Financial Time Series Analysis, et al.
- Dissertation: *Extreme gradient boosting Model (XGBoost) and its Application to Personal Credit Evaluation.*

Zhengzhou University, School of Mathematics and Statistics Zhengzhou, Henan, China

B.S. in Mathematics and Applied Mathematics 09/2014-07/2018

- Related courses: Advanced Algebra (I, II), Mathematical Analysis (I, II, III), Point Set Topology, Functional Analysis, Functions of Complex Variables, Functions of Real Variables, Modern Algebra, Differential Geometry, Computational Methods, Probability Theory, Mathematical Experiments (Mathematica), Statistics and Software Bag (R), et al.
- Dissertation: *Small and Medium Enterprises Financing based on Asset Securitization.*

PUBLICATIONS

Yumeng Cao, Lindsay Salvati, Kathryn L. Lunetta, Joanne M. Murabito, Jiachen Chen, Ahmed A.Y. Ragab, Margaret F. Doyle. *Circulating Immune Cells are Associated with Brain MRI Measurements in the Framingham Offspring*. 2023 In preparation.

Hua Li, **Yumeng Cao**, et al. (2020) *XGBoost Model and its Application to Personal Credit Evaluation.* IEEE Intelligent Systems (SCI Q2), vol.35, no.3, pp.52-61, 2020.

Wenyu Qiu, Siwen Li, **Yumeng Cao**, et al. (2019) *XGBoost Credit Evaluation Ensemble Model with Self-Contained Shunt.* The 5th International Conference on Big Data and Information Analytics, pp.59-65, 2019.

CONFERENCE PRESENTATION

(*Poster*) **Yumeng Cao**, Margaret F. Doyle, Jiachen Chen, Ahmed A.Y. Ragab, Joanne M. Murabito, Kathryn L. Lunetta. Immune Cells are Associated with Brain MRI Measurements in the Framingham Offspring, Geological Society of America Conference. Tampa, Florida, November 8 to 12, 2023.

RESEARCH EXPERIENCE

2023- Now **Framingham Immune Cell and Brain Aging Project**

Research Assistant in Dr. Lunetta's Lab
Supervisor: Kathryn L. Lunetta

- Using linear mixed effect regressions to assess potential associations between immune cell phenotypes with neuropsychological cognitive domain scores and brain MRI measurements, respectively;
- Using Cox proportional hazards regressions to explore the impact of immune cell phenotypes on the hazard of the dementia outcome and Alzheimer's disease outcome;
- Stratification Analyses based on the CMV status, sex, age, and APOE;
- Adjusting p-values for multiple correlated tests, accounting for the correlations among immune cell phenotypes.

2020-2021 **Machine Learning based on Fully Homomorphic Encryption and its Application in Risk Management**

Research Assistant in Henan Key Lab of Financial Engineering at Zhengzhou University

Supervisor: Hua Li

- Established a feasible two-party interaction model based on blockchain;
- Applied the secure logistic regression based on homomorphic encryption, and the secure extreme gradient boosting based on federated learning, on small amounts of encrypted test data based on blockchain in teamwork.

2019-2020 **The Theory and Application of Big Data on Credit Evaluation Model** (Master's Graduation Dissertation)

Research Assistant in Henan Key Lab of Financial Engineering at Zhengzhou University
Supervisor: Hua Li

- Data mining and processing of the unstructured loan data set (from Lending Club Platform);
- Carried out feature engineering using both statistical models and machine learning for imputation, normalization, batch removal and feature selection; Then built several credit evaluation models based on logistic regression and machine learning algorithms (Decision Tree, Random Forests, GBDT, and XGBoost) for training and testing;
- Evaluated results using various evaluation methods such as accuracy rate, KS, AUC and Kappa coefficient;

2019 Summer **Classification of Unbalanced Dataset in Credit Evaluation** (Cooperated with one local bank)

Research Assistant in Henan Key Lab of Financial Engineering at Zhengzhou University
Supervisor: Hua Li

- Organized, analyzed, and processed the historical personal loan data of the bank;
- Carried out feature engineering procedure, and then used grid search method for parameter optimization; Compared the methods of under-sampling, random repeat sampling and the Synthetic Minority Oversampling Technique (SMOTE) algorithms for extremely unbalanced data (default loan sample only accounts for 0.03%).

2018-2019 **Dynamic Fractal Factor Copula Method and its Application in CDO Pricing** (supported by Natural Science Foundation of China (No.11501523))

Research Assistant in Henan Key Lab of Financial Engineering at Zhengzhou University
Supervisor: Hua Li

- Established factor copula models based on different distribution for Collateralized Debt Obligation Pricing;
- Utilized Quasi-Monte Carlo simulation to realize the pricing process, and compared with the traditional Monte Carlo simulation method; Revised the algorithm independently and finished the MATLAB code with teammates;
- Participated in writing one textbook for graduate students named *Asset Securitization and its Pricing Method (taking Collateralized Debt Obligation as an example).*

2017-2018 **Small and Medium Enterprises Financing based on Asset Securitization** (Bachelor's Graduation Dissertation)

Research Assistant in Henan Key Lab of Financial Engineering at Zhengzhou University
Supervisor: Hua Li

- Analyzed the financing situation and countermeasures of small and medium enterprises in China and Europe;
- Established a pricing model for small and medium enterprises securities based on asset securitization and single factor copula model; Utilized Monte Carlo simulation to realize the pricing process and revised the algorithm and finished empirical analysis with MATLAB.

HONORS AND AWARDS

2018	**Meritorious Winner** of the Mathematical Contest in Modeling (Approximately Top 8% of teams)	International
2018	**Postgraduate Scholarship**	Provincial
2017	**Second Prize** of Higher Education Press Cup Mathematics Modeling Competition (Top 4% of teams)	National
2017-2018	**Third-Class Scholarship for Excellent Student** (Awarded to Top 15%)	University
2016-2017	**Third-Class Scholarship for Excellent Student** (Awarded to Top 15%)	University
2015-2016	**Outstanding Student Leader**	University

SKILLS

Fluent in English and Native in Chinese;
Familiar with Python, SAS, R, MATLAB, Mathematica, LaTeX, Citespace, bash;
Familiar with decision tree-based machine learning algorithms.