HANBIN LIU

Southern University of Science and Technology, 1088 Xueyuan Avenue, Shenzhen 518055, P.R. China

Homepage: `https://hanbin-liu.github.io` | Email: 12332887@mail.sustech.edu.cn

EDUCATION

Southern University of Science and Technology, China *2019.9-2023.6*
B.Sc., Statistics (Supervisor: Prof. Guoliang Tian)

- GPA: 3.88/4.00 (rank: 2/45)
- Grades (Major Courses): Advanced Linear Algebra II (H) (100), Probability Theory (100), Mathematical Statistics (100), Computational Statistics (100), Statistical Linear Models (A$^+$), Ordinary Differential Equations A (A$^+$), Time Series Analysis (A$^+$), Real Analysis (A$^+$)
- Selected to *The Honor Class* of the Department of Mathematics
- Selected to *The Elite Program* of the Department of Statistics and Data Science

Southern University of Science and Technology, China *2023.9-Present*
M.S., Math (Supervisor: Prof. Bingyi Jing)

EXPERIENCE

Southern University of Science and Technology
Large language models and their applications *2023.9-Present*

- Learning knowledge about frontier large language models and their application in different kinds areas, such as diffusion model for model reduction in molecular dynamics simulations, GPT-based LLMs for stock selection and prediction, GPT-style large time series models.

Southern University of Science and Technology **Research Intern**
Code reproduction of online debiased lasso for streaming data *2023.3-2023.6*

- Explored the lasso problem in online learning models for handling streaming data. Implemented the algorithm for online debiased lasso for streaming data. [code][slides]
- Pointed out some errors in the [paper] and contacted the author.

The Chinese University of Hong Kong, Shenzhen **Research Intern**
Missing data imputation and clustering via online high rank matrix completion *2022.1-2022.3*

- Conducted missing data imputation and clustering missions on the real medical data from some hospital in Shenzhen, which are about atherosclerosis.
- Evaluated the performance of transductive learning on the medical data using online high rank matrix completion method. Performed large-scale subspace clustering via k-factorization and compared with other methods.

The Chinese University of Hong Kong, Shenzhen **Research Intern**
Kernel-based selective sampling-based scalable sparse subspace clustering *2021.7-2021.9*

- Investigated the development from sparse subspace clustering (SSC) to selective sampling-based scalable sparse subspace clustering (S^{5}C). Derived the kernel version of S^{5}C for non-linear cases (KS^{5}C).
- Compared the alternating direction method of multipliers (ADMM) and the cyclic coordinate descent algorithm in KS^{5}C and lolved lasso problems in KS^{5}C by using the cyclic coordinate descent algorithm.
- Implemented the code for KS^{5}C and tested it. Compared with S^{5}C, there is an improvement, but is not significant, so no paper has been published. [code]

North Carolina State University **Remote Program**
The pandemic and social mobilization: agenda-building in the networked public sphere *2021.7-2021.8*

- Analyzed the corpus using a quantitative content analysis method, where the corpus was Weibo posts related to the debate on COVID-19. Created a code sheet and determined the content by the code sheet. [poster]
- Acknowledgment from "The Pandemic and Social Mobilization: Agenda-Building in the Networked Public Sphere"(ICA 2022).

PROJECTS

Statistical Analysis of Shanghai homestays based on Airbnb platform *2022.4-2022.6*

- Used SAS to conduct the explanatory data analysis. Housing prices prediction was also made.

Statistical Analysis of Short Video with R [code] [report] [slides] *2021.10-2021.12*

- Constructed a linear model to perform regression analysis on the number of likes of short videos. Conducted permutation test and bootstrapping to verify.

- Established clustering and generative models based on the number of likes, comments, and shares, as well as a classification model using type as the label.

Data Mining Applications in DC Crime [code] [report]   *2021.4-2021.5*
- Analyzed the crimes over geography and time by data preprocessing and exploratory data analysis.
- Classified and clustered the geography by the crime events. Models included decision tree, KNN, random forest, AdaBoost, GBDT, K-means, and DBSCAN.

Self-contained Report on Discrete Mathematics [report]   *2020.5-2020.6*
- Stated the methods of solving linear recurrence relations in the language of linear algebra.

AWARDS

Grants and Scholarships

China National Encouragement Scholarship	*National - 2022.11*
China National Encouragement Scholarship	*National - 2021.11*
China National Encouragement Scholarship	*National - 2020.11*
Freshman Scholarship at SUSTech	*School-level - 2019.11*

Academic Competitions

The Chinese Mathematics Competitions	*Provincial, First Prize - 2020.12*
China Undergraduate Mathematical Contest in Modeling	*Provincial, Winning Prize - 2020.11*
National Colleges Mathematics Ability Challenge, Finals	*National, Third Prize - 2020.7*
National Colleges Mathematics Ability Challenge, Preliminary	*National, Second Prize - 2020.6*
National Math Proficiency Competition for Middle School Students	*Provincial, First Prize - 2019.1*
Chinese High School Mathematics League	*Provincial, Third Prize - 2018.9*
National Mathematical Contest in Modeling for Middle School Students	*Provincial, Third Prize - 2017.9*

SKILLS AND LANGUAGE

Programming ability
- Python, R, MATLAB, Java, LaTeX.

Machine learning & Large Language Models
- Knowledge about statistical machine learning models, especially the variations of lasso, clustering algorithms, conductive learning and online learning models.
- Knowledge about deep learning models and Large Language Models (LLMs).

Language ability
- English (fluent): TOEFL (93), GRE (320)

TEACHING AND SEMINAR

MA 304 Multivariate Statistical Analysis (Teaching Assistant)	*2024.3-Present*
MA 204 Mathematical Statistics (Teaching Assistant)	*2024.3-Present*
Large Language Models (Seminar)	*2023.9-Present*
Biostatistics and Computational Statistics (Seminar)	*2021.3-2021.6*

ADDITIONAL INFORMATION

Member of the SUSTech Research project at the Center for Higher Education Research
Participated in the process of writing textbook *Mathematical Statistics*. Tutored other students in statistics.

Volunteer activities and academic activities
Vonlunteered in 2021 Shenzhen International Conference on Frontiers of Statistics and Data Science

President of the philosophy community at SUSTech
As the founder and president of the first philosophy community, I organized philosophy reading club and seminars multiple times.