

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
M.S., Math (Supervisor: Prof. Bingyi Jing)

2023.9-2025.6 (Expected)

Southern University of Science and Technology
B.Sc., Statistics

2019.9-2023.6

- 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

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 applications in different kinds of areas, such as GPT-based LLMs for stock selection and prediction, GPT-style large time series models.

Southern University of Science and Technology

Code reproduction of online debiased lasso for streaming data

Research Intern

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]
- Found some errors in the original [paper] while reproducing the code and the experiment, and subsequently contacted the author.

The Chinese University of Hong Kong, Shenzhen

Missing data imputation and clustering via online high rank matrix completion

Research Intern

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

Kernel-based selective sampling-based scalable sparse subspace clustering

Research Intern

2021.7-2021.9

- Investigated the development from sparse subspace clustering (SSC) to selective sampling-based scalable sparse subspace clustering (S⁵C). Derived the kernel version of S⁵C for non-linear cases (KS⁵C).
- Compared the alternating direction method of multipliers (ADMM) and the cyclic coordinate descent algorithm in KS⁵C and solved lasso problems in KS⁵C by using the cyclic coordinate descent algorithm.
- Implemented the code for KS⁵C and tested it. Compared with S⁵C, there is an improvement, but is not significant, so no paper has been published. [code]

PROJECTS

Statistical Analysis of Short Video with R [report] [slides]

2021.10-2021.12

- Constructed a linear model to perform regression analysis on the number of likes of short videos. Polished the regression model by using different kinds of skills (stepwise, take log, multicollinearity, etc.) and considering the real-world scenario. Conducted permutation test and bootstrapping to verify.
- Established clustering and generative models (Gaussian mixture model) based on the number of likes, comments, and shares, as well as a classification model using type as the label.

COVID-19 Hot issues and agenda building research [poster]

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.
- This is an undergraduate research project at North Carolina State University. The work was presented at the 72th Annual International Communication Association (ICA).

Data Mining Applications in DC Crime [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, K-means, DBSCAN, etc.

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. This project was completed during my sophomore year, which was not such an easy task for a sophomore. The professor gave the report high praise.

AWARDS

Grants and Scholarships

<i>China National Encouragement Scholarship</i>	<i>National - 2022.11</i>
<i>China National Encouragement Scholarship</i>	<i>National - 2021.11</i>
<i>China National Encouragement Scholarship</i>	<i>National - 2020.11</i>
<i>Freshman Scholarship at SUSTech</i>	<i>School-level - 2019.11</i>

Academic Competitions

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

SKILLS AND LANGUAGE

Programming ability

- Python, R, MATLAB, Java, L^AT_EX.

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)	<i>2024.3-Present</i>
MA 204 Mathematical Statistics (Teaching Assistant)	<i>2024.3-Present</i>
Large Language Models (Seminar)	<i>2023.9-Present</i>
Biostatistics and Computational Statistics (Seminar)	<i>2021.3-2021.6</i>

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

Volunteered 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.