Hanbin Liu

Southern University of Science and Technology, 1088 Xueyuan Avenue, Shenzhen 518055, P.R. China
Homepage: https://hanbin-liu.github.io | Email: liuhb2019@mail.sustech.edu.cn

Education

Southern University of Science and Technology, China *Sept. 2019 - June 2023 (Expected)*
B.Sc., Statistics

- GPA: 3.89/4.00 (rank: 1/48)
- 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^+)

Research Interests

Hanbin Liu's research interests lie in computational statistics, biostatistics, and the intersection of statistical machine learning and biomedical, especially some applications of statistical methods and machine learning models in biomedical research, large-scale genomic data, health and clinical data.
Some subtopics:

- **Computational Statistics**: simulation, Markov chain Monte Carlo
- **Biostatistics**: statistical genetics, bioinformatics
- **Statistical Machine Learning**: data mining, deep learning

Research Experience

Prof. Jicong Fan's Group *July 2021 - Sept. 2021*
Research Intern *CUHK-Shenzhen, School of Data Science*

- Learned sparse subspace clustering (SSC) and selective sampling-based scalable sparse subspace clustering (S^5C). Derived the kernel version of S^5C for non-linear cases (KS^5C).
- Learned the alternating direction method of multipliers (ADMM) and the cyclic coordinate descent algorithm. Solved LASSO problems in KS^5C by using the cyclic coordinate descent algorithm.
- Tested KS^5C on datasets COIL-20, COIL-100, GTSRB, YaleBCrop025, and MNIST. [code]

Prof. Alice Cheng's Group *July 2021 - Aug. 2021*
Undergraduate Research Program (Remote) *NC State, Department of Communication*

- 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]

Seminar

Biostatistics and Computational Statistics Seminar *Fall 2021*
Advisor: Prof. Guo-Liang Tian

- Studied biostatistics with R. Studied the EM algorithm and its derivatives, QLB algorithm, and De Pierro algorithm.

Project

Statistical Analysis of Short Video with R [code] [report] [slides] *Fall 2021*

- 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] *Spring 2021*

- 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] *Fall 2020*

- Stated the methods of solving linear recurrence relations in the language of linear algebra.

Selected Awards and Honors

- China National Encouragement Scholarship *Nov. 2021*
- The 12th Chinese Mathematics Competitions (Provincial First Prize) *Dec. 2020*
- China National Encouragement Scholarship *Nov. 2020*

Skills and Languages

- R, Python, MATLAB, Java, LaTeX
- English (fluent), Mandarin (native)

Additional Information

- Volunteer: 2021 Shenzhen International Conference on Frontiers of Statistics and Data Science
- Coursework: `https://github.com/Hanbin-Liu/sustech-assign`