Haitian Zhong

Mathematics and applied mathematics
Cuiying Honors College
Lanzhou University
222 Tianshui South Road, Lanzhou, Gansu, China

☎ +86-15996295520
✉ haitian.zhong@cripac.ia.ac.cn
✉ zhonght20@lzu.edu.cn
⌸ GitHub
⊕ Homepage

EDUCATION

- **The Institute of Automation of the Chinese Academy of Sciences (CASIA)** *From Sept. 2024(Expected)*
  The Center for Research on Intelligent Perception and Computing (CRIPAC) Beijing, China

  – Research interests: Large Language Models(MLLM Editing and Reasoning), Numerical Analysis(Computational PDEs), Machine Learning(Generative Models and Theory) and AI for Science(Bioinformatics)
  – Supervisor: Prof. Tieniu Tan
  – Co-supervisor: Prof. Qiang Liu and Prof. Shu Wu

- **Lanzhou University** *Sept. 2020 - Present*
  Mathematics and applied mathematics (China's Top-notch Undergraduate Training Program 2.0) Lanzhou, China

  – GPA: 91.70/100
  – Ranking: 2/152
  – Language: CET-4: 656; CET-6: 601(Oral grading: B+); TOEFL: 103(in 2017[1])
  – Related Courses: Mathematical Analysis(95), Advanced Algebra(96), Probability Theory(95), Real Analysis(98), Numerical Analysis(97), Abstarct Algebra(98), Mathematical Statistics(98), Graph Theory(95), Operational Research(96), etc.

- **Nanjing Foreign Languages School** *Sept. 2014 - Jun. 2020*
  Top Class of Science Nanjing, China

  – Selected Awards: (Senior) High School Mathematics Competition, Provincial First Prize; Zhou Enlai's Merit Scholarship; Nanjing Merit Student; Merit Student of NFLS (6-star).

SELECTED AWARDS

- **S.-T. Yau College Student Mathematics Contest(Applied and Computational Maths)**,Excellence award *Jul. 2022*
- **S.-T. Yau College Student Mathematics Contest(Analysis and Differential Equations)**,Excellence award *Jul. 2022*
- **The 14th Chinese Mathematics Competition**,Provincial First Prize(Top 1%, Promoted to National Final) *Apr. 2023*
- **The 13th Chinese Mathematics Competition**,National Second Prize(Top 1%) *Mar. 2023*
- **University Merit Scholarship**,First Class(Top 1%) *Nov. 2022*
- **"FLTRP-ETIC Cup" English Public Speaking Contest**,First Prize *Sept. 2022*
- **National English Competition for College Students**,National First Prize(Top 1%) *May 2022*
- **University Merit Scholarship**,First Class(Top 1%) *Nov. 2021*

PROJECTS

- **Knowledge editing for Multimodel Large Language Models** *Oct. 2023-Present*
  Advised by Prof. Qiang Liu and Shu Wu during the internship at CRIPAC, CASIA Beijing, China

  – Abstract: Despite the ability to train capable Large Language Models(LLMs), the methodology for maintaining their relevancy and rectifying errors remains elusive. To this end, the past few years have witnessed a surge in techniques for editing LLMs, the objective of which is to efficiently alter the behavior of LLMs within a specific domain without negatively impacting performance across other inputs. Following the trends, we wish to implement knowledge editing on Mulitimodel Large Language Models(MLLMs). We will build a new benchmark dataset to evaluate the editing performance on variable metrics including Portability. What's more, we wish to carry sequential editing on MLLMs as well, making lifelong targeted edits while avoiding expensive retraining.

- **PTransIPs:Identification of SARS-CoV-2 phosphorylation sites based on pretrained embedding and transformer** *Nov. 2022-Present*
  Collaborate with Ziyang Xu(LZU), Xueying Wang*(CityU), Tianchi Lu*(CityU) Lanzhou, China

  – Abstract: Identification of phosphorylation sites is an important step for understanding the molecular mechanisms of SARS-CoV-2 infection and the changes within the host cells pathways. In this study, we present PTransIPs, a deep-learning architecture to identify phosphorylation sites in host cells infected with SARS-CoV-2. PTransIPs utilizes two protein pretrained models and the popular transformer structure to make the final prediction, and we also add the transductive information maximization (TIM) loss to better evaluate the error.

[1]Acquired in high school

- Performance: After comparing it with five existing phosphorylation site prediction tools using independent test data and adopting the reported independent test performance for comparison, we note that PTransIPs has superior predictive ability compared with existing five predictorsachieving the best performance in all five model evaluation metrics (ACC, SEN, SPEC, MCC, AUC) for both S/T and Y sites. Notably, we found that the addition of structural information does not greatly improve the prediction for Y sites, and this should be related to the small amount of data for Y sites. This implies that structural information might not effectively enhance the model performance on small sample dataset. We hope that PTransIPs will aid in deepening the understanding of SARS-CoV-2 phosphorylation sites and look forward to enhancing PTransIPs in the f uture to become a more powerful tool for the scientific community.
- See our paper at arXiv:2308.05115, codes at Github:PTranIPs

- **Nodal discontinuous Galerkin Method for Chemotaxis - Shallow Water System** *Jul. 2022-Present*
  Advised by Prof. Weihua Deng(LZU) Lanzhou, China
  - Abstract: The chemotaxis shallow water system consists of several hyperbolic equations which describe the behavior of the velocity of the flow and the interactions between the concentration of the chemicals and the cells dissolved within. The system is derived from the incompressible Navier-Stokes equation, in the case where the horizotal length scale is appreciably larger than the vertical one.It's clear that for a bounded domain in $\mathbb{R}^3$ the system acquires a unique stable global solution. In this work we combine the characteristic method and the nodal discontinuous Galerkin method with carefully constructing numerical fluxes. The nonlinear stability is desired for the proposed symmetric variational formulation. Moreover, for general triangulations the priori estimates for the L^2-norm of the errors in both velocity and concentrations will be derived. Finally, some numerical experiments are performed to verify theoretical results.

Exchange Experiences

- **The 9th International Forum on Statistics** *Jul. 2023*
  Mingde Building, Renmin University of China Beijing, China

  - Plenary Talks: Fast Distributed Principal Component Analysis of Large-Scale Federated Data (Xihong Lin, Harvard University); On Dynamics-Informed Blending of Machine Learning and Microeconomics Speaker (Michael I. Jordan, University of California, Berkeley); Optimal Nonparametric Testing of Missing Completely at Random, and Its Connections to Compatibility (Richard J. Samworth, University of Cambridge).
  - Distinguished Session: Generating Robust Evidence with Multi-institutional EHR Data (Tianxi Cai, Harvard University); Genetic Studies of Human Brain Imaging Data (Heping Zhang, Yale University).
  - Invited Sessions: 2D-Shapley: A Framework for Fragmented Data Valuation (Xiangyu Chang,Xi'an Jiaotong University); Consistent Selection of the Number of Groups in Panel Models via Sample-Splitting (Xuening Zhu, Fudan University); Efficient, Stable, and Analytic Differentiation of the Sinkhorn Loss (Yixuan Qiu, Shanghai University of Finance and Economics); Peer-Label Assisted Hierarchical Text Classification (Feifei Wang, Renmin University of China).

- **Summer Semester, 2021** *Jul. 2021*
  Cuiying Honors College, Lanzhou University Lanzhou, China

  - Coursework: Multi-scale Models, Algorithm and Analysis(Pingbing Ming, Chinese Academy of Sciences); Integrating Data and Dynamics in scRNA-seq Data Analysis(Tiejun Li, Peking University); Distance-regular graphs(Jacobus Hendricus Koolen, University of Science and Technology of China).

Integrated Skills

My skills consist of Mathematics knowlegde, Computer Programming and excellent language ability. Nevertheless, I am a fast learner of new tools and a fanatic lover of self-learning.

- **Mathematics**: Optimzation, Analysis, PDE, Numerical Analysis, Statistics

- **Programming Languages**: Python, R, C/C++, Mathematica, MATLAB, LaTeX

- **Artificial Intelligence**: PyTorch

- **English**: Very fluent in oral English; Proficient in English writing and reading English papers