

GE LIU

+1 617 955 0325 ◊ geliu@csail.mit.edu ◊ <https://people.csail.mit.edu/geliu/>

[Google Scholar]

EDUCATION

Massachusetts Institute of Technology August 2015 - August 2020
Ph.D. in Computer Science, advised by Prof. David Gifford, CSAIL, GPA: 5.0/5.0

Tsinghua University, China August 2011 - July 2015
Bachelor of Engineering in Electronic Engineering, GPA: 3.78/4.00

HONORS AND AWARDS

George M. Sprowls Ph.D. Thesis Award in AI and Decision-Making 2021
(awarded to 2 PhD students in MIT EECS, and I am the **only female recipient**)
Finalist in The Data Open Datathon global championship in 2019, and Boston-regional top 3 in 2018

David S. Y. Wong Fellowship at MIT 2016
Friendship of Tsinghua-Samsung Scholarship. 2013
Outstanding Social Work Scholarship. 2012
2nd prize in Chinese Physics Olympiad for High school, Beijing 2011
Massachusetts Math Olympiad Level One, Finalist, top 25 of Massachusetts State. 2008

PROFESSIONAL EXPERIENCE

Adjunct Professor (Incoming Assistant Professor), CS Department, UIUC Jun 1, 2023 - now
I will be joining the Computer Science department at University of Illinois Urbana-Champaign as an Assistant Professor in 2024, leading a research group in machine learning and computational biology.

Senior Applied Scientist, AWS AI Lab, Palo Alto Sep 14, 2020 - now
I am working with AWS AI Labs on large-scale deep sequential modeling, personalized recommender systems, large language model and time series modeling. I led two high-impact projects on audience modeling and turned them into successfully AI services on AWS.

Research Intern, Google Brain, Mountain View May 27, 2019 - Aug 30, 2019
I was working with Mixel(SIR) team at Google Brain on inventing novel data-efficient training approach for Reinforcement Learning with Adaptive Behavior Policy Sharing.

Software Engineering Intern (Ph.D.), Google Brain, Mountain View Jun 11, 2018 - Aug 31, 2018
I was working with medical audio scribe modeling team at Google Brain on active learning for natural language processing tasks in health-care domain.

Undergraduate Research Assistant, advised by professor Jianyang Zeng Sept 2014 - Apr 2015
Machine Learning and CompBio Group, IIIS, Tsinghua University

Visiting Student Researcher, advised by professor Robert F. Murphy Jul 2014 - Sept 2014
Murphy Lab, Lane Center for Computational Biology, Carnegie Mellon University

RESEARCH HIGHLIGHTS

Computational peptide vaccine design and evaluation

- I developed end-to-end AI-driven vaccine development framework (OptiVax) for evaluating T-cell peptide vaccine efficiency and designing optimal peptide vaccines[Cell systems 2020], unveiled insufficient population coverage of COVID-19 mRNA vaccines among under-represented ethnic groups, and proposed augmentation [Cell systems 2021]. Formulated vaccine design as a novel combinatorial optimization problem and proposed effective algorithms[ICLR 2022a]. The proposed vaccine was able to prevent mortality in transgenic mice infected with SARS-CoV-2 Beta with 5.5 times more CD8+ T cell infiltration than Pfizer vaccine [Frontiers in Immunology 2023]. My research fueled life science startup ThinkTherapeutics which uses the technique to design a new cancer treatment

that is in an early stage clinical trial. Our research has been featured by major media such as Boston Globe, MIT News, Amsterdam News, etc.

Deep-learning-driven *de novo* therapeutic antibody design

- My research develops first-ever ML-driven design framework that revolutionized iterative screening-based molecule design and generated monoclonal antibodies with experimentally verified better affinity [Bioinformatics 2019] and specificity [Cell reports methods 2022] without the need of structural information with better efficiency.

Explainable deep learning for functional genomics

- Established the very first standard guidelines for architecture and benchmark design for deep genomics sequence modeling [ISMB 2016, Bioinformatics 2016], and developed tools to interpret deep feature interactions learned by deep genomics neural network [BMC Bioinformatics 2019].

Uncertainty-aware and data-efficient ML for iterative learning under limited data

- I worked on methods for improving the quality of model uncertainty for deep ensembles and showed its effectiveness in OOD prediction and Bayesian optimization [AAAI 2020]. I am also interested in utilizing model uncertainty for exploration in antibody optimization [Bioinformatics 2019] and data-efficient Reinforcement Learning [NeurIPS DeepRL workshop 2019], and developing novel active learning strategies such as ICAL [under review].

Advancing techniques for sequential modeling

- Exploring the use of TPP-based intensity modeling for user targeting and joint user-item recommendation which optimizes for both relevance and diversity [ICLR 2022b]. Utilizing graph information in Transformer-based recommender systems for better user representation learning [NeurIPS workshop].

SELECTED PUBLICATIONS AND PATENTS

Journal Papers

1. Brandon Carter, Pinghan Huang, **Ge Liu**, Yuejin Liang, Paulo JC Lin, Bi-Hung Peng, Lindsay McKay, Alexander Dimitrakakis, Jason Hsu, Vivian Tat, Panatda Saenkham-Huntsinger, Jinjin Chen, Clarety Kaseke, Gaurav D Gaiha, Qiaobing Xu, Anthony Griffiths, Ying K Tam, Chien-Te K Tseng, David K Gifford. “A pan-variant mRNA-LNP T cell vaccine protects HLA transgenic mice from mortality after infection with SARS-CoV-2 Beta.” *Frontiers in Immunology*, 2023
2. Sachit Dinesh Saksena, **Ge Liu**, Christine Banholzer, Geraldine Horny, Stefan Ewert, David K Gifford. “Computational counterselection identifies nonspecific therapeutic biologic candidates” *Cell reports methods*, 2022
3. **Ge Liu**, Brandon Carter, and David K. Gifford. “Predicted Cellular Immunity Population Coverage Gaps for SARS-CoV-2 Subunit Vaccines and their Augmentation by Compact Peptide Sets.” *Cell Systems*, 2021
4. **Ge Liu**, Brandon Carter, Trenton Bricken, Siddhartha Jain, Mathias Viard, Mary Carrington, and David K. Gifford. “Computationally Optimized SARS-CoV-2 MHC Class I and II Vaccine Formulations Predicted to Target Human Haplotype Distributions” *Cell Systems*, 2020
5. **Ge Liu**, Haoyang Zeng, Jonas Mueller, Brandon Carter, Ziheng Wang, Jonas Schilz, Geraldine Horny, Michael E Birnbaum, Stefan Ewert, David K Gifford. “Antibody Complementarity Determining Region Design Using High-Capacity Machine Learning” *Bioinformatics*, 2019
6. **Ge Liu**, Haoyang Zeng, David K Gifford . “Visualizing Complex Feature Interactions and Feature Sharing in Genomic Deep Neural Networks” *BMC bioinformatics 20.1 (2019): 401*.
7. Haoyang Zeng, Matthew D Edwards, **Ge Liu**, David K Gifford. “Convolutional neural network architectures for predicting DNA-protein binding.” *Intelligent Systems for Molecular Biology (ISMB)*, 2016, and *Bioinformatics*, 32(12):i121-i127, 2016.
8. **Ge Liu**. “Beyond predictive modeling: new computational aspects for deep learning based biological applications” *Ph.D. Thesis, MIT*, 2020, George M. Sprowls **Ph.D. Thesis Award** in AI and Decision-Making

Conference and Workshop Papers

9. **Ge Liu**, Alexander Dimitrakakis, Brandon Carter, and David K. Gifford. “Maximum n-times Coverage for Vaccine Design” *Proceedings of the 10th International Conference on Learning Representations (ICLR)*, 2022
10. Yifei Ma, **Ge Liu**, Anoop Deoras. “Bridging Recommendation and Marketing via Recurrent Intensity Modeling” *Proceedings of the 10th International Conference on Learning Representations (ICLR)*, 2022
11. Zeren Shui, **Ge Liu**, Anoop Deoras, George Karypis “Sequence-Graph Duality: Unifying User Modeling with Self-Attention for Sequential Recommendation.” *NeurIPS Workshop on New Frontiers in Graph Learning*, 2022
12. **Ge Liu***, Siddhartha Jain*, Jonas Mueller, and David Gifford. “Maximizing Overall Diversity for Improved Uncertainty Estimates in Deep Ensembles” *Proceedings of the 34th Conference on Artificial Intelligence(AAAI)* 2020
13. Yifei Ma, **Ge Liu**, Anoop Deoras. “Recurrent Intensity Modeling for User Recommendation and Online Matching” *Time Series Workshop at ICML*, 2021
14. **Ge Liu**, Heng-tze Cheng, Rui Wu, Jing Wang, Jayden Ooi, Sibon Li, Ang Li, Lihong Li, Craig Boutilier, Ed Chi. “Data Efficient Training for Reinforcement Learning with Adaptive Behavior Policy Sharing.” work done during internship at Google Brain (in collaboration with DeepMind), *NeurIPS Deep Reinforcement Learning Workshop and Optimization Foundations of RL workshop*, 2019
15. **Ge Liu**, David K Gifford. “Visualizing Feature Maps in Deep Neural Networks using DeepResolve-A Genomics Case Study” *Workshop on Visualization for Deep Learning at ICML*, 2017.

Under Review

16. **Ge Liu**, Shima Alizadeh, Aniruddha Bhargava, Karthick Gopalswamy, Lalit K Jain, Branislav Kveton “Pessimistic Multi-Objective Optimization.” *Under review at NeurIPS 2023*
17. Siddhartha Jain, **Ge Liu**, David K Gifford. “Information Condensation Active Learning.” *Under review*
18. **Ge Liu**, Haoyang Zeng, David K Gifford. “Machine learning based antibody design” *US20190065677A1, pending*, 2019.

ACADEMIC SERVICES AND TEACHING

Program Committee member: ICML 2020/2021/2022(**Outstanding reviewer award**), NeurIPS 2021/2022/2023, ICLR 2022/2023, AISTATS 2021/2022/2023, UAI 2021/2022/2023, IJCAI 2021/2022/2023, RECOMB, ICMLA

Reviewer for journals: Bioinformatics, PLOS Computational Biology, IEEE/ACM Transactions on Computational Biology and Bioinformatics, BioData Mining, TMLR, Machine Learning

Workshop and tutorial organizer: Tutorial on Deep User Targeting, AMLC 2021; Uncertainty-aware Deep Learning Workshop, AMLC 2022

Teaching Assistant: MIT 6.874/6.802/20.390/20.490/HST.506 (2019 Spring) - “Computational Systems Biology: Deep Learning in the Life Sciences”

GRANT WRITING EXPERIENCE

I helped with the writing that lead to the following funding

NIH R01CA218094: Deep-learning based antibody design using high-throughput affinity testing of synthetic sequences

C3.ai DTI Research Award Program: Machine Learning Based Vaccine Design and HLA Based Risk Prediction for Viral Infections

GCP research credits program: Data Efficient Training for Reinforcement Learning with Adaptive Behavior Policy Sharing

INVITED TALKS

- “Accelerating Biomedical Discovery and Therapeutic Biologic Design With ML”, *Dept.CS, UT Austin*, Apr 2023
- “Accelerating Biomedical Discovery and Therapeutic Biologic Design With ML”, *Dept.CS, U Chicago*, Mar 2023
- “Accelerating Biomedical Discovery and Therapeutic Biologic Design With ML”, *NYU Courant Institute*, Mar 2023
- “Accelerating Biomedical Discovery and Therapeutic Biologic Design With ML”, *Dept.CS, UIUC*, Feb 2023
- “Deep User Targeting Systems”, Tutorial, *Amazon AMLC 2021*, Sep 2021
- “Maximizing Overall Diversity for Improved Uncertainty Estimates in Deep Ensembles”, *Fidelity AI*, Aug 2020
- “Information Condensing Active Learning”, *Google ML Research Talk Series*, May 2020

MENTORING EXPERIENCES

- Ziheng Wang** (Undergrad at MIT, now a CS PhD at Stanford)
- Trenton Bricken** (Undergrad at Duke, now a PhD at Harvard)
- Alexander Dimitrakakis** (Undergrad and Master at MIT, incoming MD student)
- Brandon Carter** (Undergrad and CS PhD at MIT, now the CTO of ThinkTherapeutics)
- Sachit Dinesh Saksena** (PhD in Computational and Systems Biology at MIT)
- Bianca Lepe** (Bioengineering PhD at MIT)
- Nathan R Hunt** (Undergrad and CS PhD at MIT)
- Zeren Shui** (CS PhD at University of Minnesota, summer intern of AWS AI Labs)

SELECTED MEDIA COVERAGE

- “Creating a versatile vaccine to take on Covid-19 in its many guises” - *MIT News*
- “Local scientists hope to create the ultimate COVID vaccine” - *Boston Globe*
- “Coronavirus vaccines may be less effective for Black and Asian recipients, MIT study suggests.” - *Yahoo News*
- “Study: adding peptides to COVID-19 vaccines improves effectiveness for under-represented populations” - *MIT CSAIL News*
- “MIT machine learning models find gaps in coverage by Moderna, Pfizer, other Warp Speed COVID-19 vaccines” - *ZDNet*
- “MIT’s machine learning designed a COVID-19 vaccine that could cover a lot more people” - *ZDNet*
- “Potential COVID-19 vaccines get a boost from machine learning” - *MIT CSAIL News*
- “COVID-19 vaccines may be less effective for racial minorities, suggest computer scientists” - *European Pharmaceutical Review*
- “VACCINE AND RACE: Report says COVID vaccine may not be as effective for Blacks” - *Amsterdam News*
- “Reports of Racial Disparities in Covid Vaccines Distort Science” - *Fairness & Accuracy In Reporting*
- “Covid-19 vaccines may be less effective for racial minorities” - *Tech Explorist*
- “MIT’s machine learning designed a COVID-19 vaccine that could cover a lot more people” - *AiBots*
- “AI Shows COVID-19 Vaccines May Be Less Effective in Racial Minorities.” - *HealthITAnalytics*
- “MIT Researchers Leverage Machine Learning for COVID Vaccine Research” - *Pure AI*
- “MIT develops machine learning model to quicken release of COVID-19 vaccine” - *TechRepublic*