Tianfan Fu

CONTACT INFORMATION

Computational Science and Engineering
Georgia Institute of Technology
Atlanta, GA, US
tfu42@gatech.edu & futianfan@gmail.com
Tel: +1 4706013173
`https://futianfan.github.io/`

RESEARCH INTERESTS

AI for Drug Discovery and Development, AI for Healthcare, AI for Science.

EDUCATION

Georgia Institute of Technology, Atlanta, US.

Ph.D. candidate, Advisor: **Jimeng Sun**, Computational Science and Engineering, August 2018 - Present.
Research topic: **AI for Drug Discovery and Development**.

Shanghai Jiao Tong University (SJTU), Shanghai, CHINA

M.S., Advisor: **Zhihua Zhang**, Computer Science and Technology, Sept 2015 - March 2018.
Research topic: Bayesian computation and inference.

Shanghai Jiao Tong University (SJTU), Shanghai, CHINA

B.E., Electronics and Electric Engineering (IEEE Honor Class), Computer Science, Sept 2011 - June 2015.
Advisor: **Kai Yu**.
Research topic: deep learning on speech recognition and speaker verification.

INDUSTRY EXPERIENCE

Research Intern, Machine Learning Group, IQVIA, Boston

Advisor: **Cao Xiao**, Summer 2021 (May - Aug) & Summer 2020 (May - Aug)
Research topic: clinical trial outcome prediction

Research Intern, Machine Learning Group, Disney Research Institute, Pittsburgh

Advisor: **Stephan Mandt**, Oct 2017 - Nov 2017.
Research topic: word/user embeddings algorithm

TEXTBOOK

Tianfan Fu, Cao Xiao, Jimeng Sun: Machine learning for drug discovery and development. Expect to complete in May 2023 by Springer. `https://ml4drug-book.github.io/`

Shenjian Zhao, Yujun Li, **Tianfan Fu**, Kai Li, Zhihua Zhang: **Chinese Translation of "Deep Learning (Goodfellow et al)"**. Sales volume: receiving **200K+** comments in `jd.com`.

PUBLICATIONS

Hanchen Wang*, **Tianfan Fu***, Yuanqi Du*, Wenhao Gao, Kexin Huang, Ziming Liu, Payal Chandak, Shengchao Liu, Peter Van Katwyk, Andreea Deac, Anima Anandkumar, Karianne Bergen, Carla P. Gomez, Shirley Ho, Pushmeet Kohli, Joan Lasenby, Jure Leskovec, Tie-Yan Liu, Arjun Manrai, Debora Marks, Bharath Ramsundar, Le Song, Jimeng Sun, Jian Tang, Petar Veličković, Max Welling, Connor Coley, Yoshua Bengio, Marinka Zitnik: Enabling Scientific Discovery with Artificial Intelligence, accepted by **Nature**, 2023.

Tianfan Fu*, Wenhao Gao*, Connor W. Coley, Jimeng Sun. Reinforced Genetic Algorithm for Structure-based Drug Design. Neural Information Processing Systems (**NeurIPS**) 2022.

Wenhao Gao*, **Tianfan Fu***, Jimeng Sun, Connor W. Coley: Sample Efficiency Matters: A Benchmark for Practical Molecular Optimization. Neural Information Processing Systems (**NeurIPS** 2022) Track on Datasets and Benchmarks.

Kexin Huang*, **Tianfan Fu***, Wenhao Gao*, Yue Zhao, Yusuf Roohani, Jure Leskovec, Connor W. Coley, Cao Xiao, Jimeng Sun, Marinka Zitnik: Artificial Intelligence Foundation for Therapeutic Science. **Nature Chemical Biology**, 2022.

Tianfan Fu, Jimeng Sun: SIPF: Sampling Method for Inverse Protein Folding. The 28th ACM **SIGKDD** Conference on Knowledge Discovery and Data Mining (2022).

Tianfan Fu, Jimeng Sun: Antibody Complementarity Determining Regions (CDRs) design using Constrained Energy Model. The 28th ACM **SIGKDD** Conference on Knowledge Discovery and Data Mining (2022).

Tianfan Fu*, Wenhao Gao*, Cao Xiao, Jacob Yasonik, Connor W. Coley, Jimeng Sun. Differentiable Scaffolding Tree for Molecular Optimization. International Conference on Learning Representation (**ICLR**), 2022.

Tianfan Fu, Kexin Huang, Cao Xiao, Lucas M. Glass, Jimeng Sun. HINT: Hierarchical Interaction Network for Clinical Trial Outcome Prediction. Cell Patterns, 2022. **cover paper of Cell Patterns**, `https://www.cell.com/patterns/issue?pii=S2666-3899(21)X0005-0`

Kexin Huang*, **Tianfan Fu***, Wenhao Gao*, Yue Zhao, Yusuf Roohani, Jure Leskovec, Connor W. Coley, Cao Xiao, Jimeng Sun, Marinka Zitnik: Therapeutics Data Commons: Machine Learning Datasets and Tasks for Drug Discovery and Development. Neural Information Processing Systems (**NeurIPS** 2021) Track on Datasets and Benchmarks.

Tianfan Fu, Cao Xiao, Lucas Glass, Jimeng Sun: MOLER: Incorporate Molecule-Level Reward to Enhance Deep Generative Model for Molecule Optimization. IEEE Transactions on Knowledge and Data Engineering (**TKDE**) 2021.

Tianfan Fu, Cao Xiao, Cheng Qian, Lucas Glass, Jimeng Sun: Probabilistic and Dynamic Molecule-Disease Interaction Modeling for Drug Discovery. The 27th ACM **SIGKDD** Conference on Knowledge Discovery and Data Mining (2021).

Tianfan Fu, Cao Xiao, Xinhao Li, Lucas Glass, Jimeng Sun: MIMOSA: Multi-constraint Molecule Sampling for Molecule Optimization. Association for the Advancement of Artificial Intelligence (**AAAI**) 2021.

Kexin Huang, **Tianfan Fu**, Lucas Glass, Marinka Zitnik, Cao Xiao, Jimeng Sun: DeepPurpose: a Deep Learning Library for Drug-Target Interaction Prediction. **Bioinformatics** 2020.

Tianfan Fu, Cao Xiao, Jimeng Sun: CORE: Automatic Molecule Optimization using Copy & Refine Strategy. Association for the Advancement of Artificial Intelligence (**AAAI**) 2020.

Tianfan Fu*, Tian Gao*, Cao Xiao, Tengfei Ma, Jimeng Sun: PEARL: Prototype Learning via Rule Learning. ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (**ACM BCB**) 2019.

Tianfan Fu*, Trong Nghia Hoang*, Cao Xiao, Jimeng Sun: DDL: Deep Dictionary Learning for Predictive Phenotyping. International Joint Conferences on Artificial Intelligence (**IJCAI**) 2019.

Tianfan Fu, Cheng Zhang, Stephan Mandt: Continuous Word Embedding Fusion via Spectral Decomposition. SIGNLL Conference on Natural Language Learning (**CoNLL**) 2018.

Tianfan Fu, Zhihua Zhang: CPSG-MCMC: Clustering-Based Preprocessing method for Stochastic Gradient MCMC. Artificial Intelligence and Statistics, (**AISTATS**) 2017.

Tianfan Fu, Luo Luo, Zhihua Zhang: Quasi-Newton Hamiltonian Monte Carlo. Conference on Uncertainty in Artificial Intelligence (**UAI**) 2016.

Wei Li, **Tianfan Fu**, Hanxu You, Jie Zhu, Ning Chen: Feature sparsity analysis for i-vector based speaker verification. **Speech Communication** 2016.

Yuan Liu, Yanmin Qian, Nanxin Chen, **Tianfan Fu**, Ya Zhang, Kai Yu: Deep feature for text-dependent speaker verification. **Speech Communication**, 2015. (**2019 EURASIP award for the best paper published in Speech Communication (2014-2017)**)

Wei Li, **Tianfan Fu**, Jie Zhu: An improved i-vector extraction algorithm for speaker verification. **EURASIP J. Audio, Speech and Music Processing** 2015.

Wei Deng, Yanmin Qian, Yuchen Fan, **Tianfan Fu**, Kai Yu: Stochastic data sweeping for fast DNN training. IEEE International Conference on Acoustics, Speech and Signal Processing (**ICASSP**) 2014.

Yuan Liu, **Tianfan Fu**, Yuchen Fan, Yanmin Qian, Kai Yu: Speaker verification with deep features. International Joint Conference on Neural Networks (**IJCNN**) 2014.

Tianfan Fu, Yanmin Qian, Yuan Liu, Kai Yu: Tandem deep features for text-dependent speaker verification (**INTERSPEECH**) 2014.

Patent

Tianfan Fu, Kexin Huang, Jimeng Sun: Automated prediction of clinical trial outcome. United States Patent Application 20230034559. `https://www.freepatentsonline.com/y2023/0034559.html`.

Yuchen Fan, Zheng Li, Liya Yang, **Tianfan Fu**, Kai Yu, Yanlong Wang: Mobile authentication system and method based on voiceprint recognition, face recognition and location service. CN103440686A. Shanghai Jiao Tong University.

Journal Reviewer

- Cell Patterns, 2022.
- Frontiers Bioengineering, 2020.
- Frontiers in Genetics, 2020.
- IEEE Journal of Biomedical and Health Informatics (JBHI), 2020.
- IEEE Transactions on Cybernetics, 2020.
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS), 2021.
- PLOS Computational Biology, 2020.

Conference Reviewer

- AAAI - AAAI Conference on Artificial Intelligence, 2017, 2018, 2021.
- CHIL - Conference on Health, Inference, and Learning, 2022.
- ICLR - International Conference on Learning Representations, 2022.
- ICML - International Conference on Machine Learning, 2021, 2022.
- IJCAI - International Joint Conference on Artificial Intelligence, 2021.
- NeurIPS - Neural Information Processing Systems, 2016, 2020, 2021, 2022.

Workshop Organizer

- NeurIPS 2021 "1st AI for Science: Mind the Gaps" Workshop
- ICML 2022 "2nd AI for Science" Workshop.
- NeurIPS 2022 "3rd AI for Science: Progress and Promises" Workshop. `https://ai4sciencecommunity.github.io/`

Teaching

- 2016 Spring Prof. Zhihua Zhang's course "Statistical Machine Learning" TA
- 2018 Spring Prof. Bo Yuan's course "Artificial Intelligence" TA

- 2019 Fall Prof. Jimeng Sun's course "Big Data Analytics for Healthcare" TA
- 2020 Spring Prof. Jimeng Sun's course "Big Data Analytics for Healthcare" TA
- 2020 Fall Prof. Jimeng Sun's course "Big Data Analytics for Healthcare" TA

References

- Prof. Jimeng Sun (Ph.D. Advisor)
  - Professor in Computer Science Department & Carle's Illinois College of Medicine, University of Illinois, Urbana-Champaign (UIUC).
  - Email: jimeng@illinois.edu
  - Interfolio email: send.Sun.1390BE0480@interfolio.com
  - Homepage: `https://sunlab.org/`
- Prof. Connor W. Coley
  - Assistant Professor in Chemical Engineering & Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT).
  - Email: coley-office@mit.edu
  - Interfolio email: send.Coley.F1EE6D2B69@interfolio.com
  - Homepage: `https://coley.mit.edu/`
- Prof. Marinka Zitnik
  - Assistant Professor of Biomedical Informatics, Harvard Medical School, Harvard University.
  - Email: marinka@hms.harvard.edu
  - Interfolio email: send.Zitnik.23A451D4C0@interfolio.com
  - Homepage: `https://dbmi.hms.harvard.edu/people/marinka-zitnik`