Tianfan Fu

CONTACT INFORMATION

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

RESEARCH INTERESTS

AI for Drug Discovery, Machine Learning for Healthcare, Natural Language Processing, Speech Processing, Bayesian Statistics.

EDUCATION

Georgia Institute of Technology, Atlanta, US.

Ph.D. candidate, Advisor: **Jimeng Sun**, Computer Science Program in Department of Computational Science and Engineering, August 2018 - Present.

Shanghai Jiao Tong University (SJTU), Shanghai, CHINA

M.S., Advisor: **Zhihua Zhang**, Computer Science and Technology, Sept 2015 - March 2018.

Shanghai Jiao Tong University (SJTU), Shanghai, CHINA

B.E., Electronics and Electric Engineering (IEEE Honor Class), Computer Science, Sept 2011 - June 2015.

RESEARCH EXPERIENCE

Research Assistant, Speech Lab, Shanghai Jiao Tong University (SJTU)

Advisor: **Kai Yu**, June 2013 - Jan 2015.
Research topic: application of deep learning on speech recognition and speaker verification.

Research Assistant, Learning and Optimization Group, Shanghai Jiao Tong University (SJTU)

Advisor: **Zhihua Zhang**, Feb 2015 - May 2017.
Research topic: Bayesian computation and inference.

Research Assistant, SunLab, Georgia Institute of Technology

Advisor: **Jimeng Sun**, Sept 2018 - Present.
Research topic: Drug Discovery and Development.

INDUSTRY EXPERIENCE

Research Intern, Machine Learning Group, IQVIA, Boston

Advisor: **Cao Xiao**, May 2020 - Aug 2020.
Research topic: clinical trial prediction

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

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

Intern, Dialogue System Group, AISPEECH, Suzhou, China

Project: Text Similarity for QA system, Feb 2018 - June 2018.

PUBLICATIONS

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, New York, NY, USA. (**Oral**)

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, Niagara Fall, NY, USA. (**Oral**)

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

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

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`. **深度学习中译版** 京东评论200K+条

Tianfan Fu, Zhihua Zhang: CPSG-MCMC: Clustering-Based Preprocessing method for Stochastic Gradient MCMC. **AISTATS** 2017: 841-850, Lauderdale, FL, USA. (**Poster**)

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

Wei Li, **Tianfan Fu**, Hanxu You, Jie Zhu, Ning Chen: Feature sparsity analysis for i-vector based speaker verification. **Speech Communication** 80: 60-70, 2016.

Yuan Liu, Yanmin Qian, Nanxin Chen, **Tianfan Fu**, Ya Zhang, Kai Yu: Deep feature for text-dependent speaker verification. **Speech Communication** 73: 1-13, 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: 18, 2015.

Wei Li, **Tianfan Fu**, Jie Zhu, Ning Chen: Sparsity Analysis and Compensation for i-Vector Based Speaker Verification. **SPECOM** 2015: 381-388.

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: 240-244.

Yuan Liu, **Tianfan Fu**, Yuchen Fan, Yanmin Qian, Kai Yu: Speaker verification with deep features. International Joint Conference on Neural Networks, **IJCNN** 2014: 747-753, Beijing, China (**Oral**)

Tianfan Fu, Yanmin Qian, Yuan Liu, Kai Yu: Tandem deep features for text-dependent speaker verification. **INTERSPEECH** 2014: 1327-1331, Singapore. (**Oral**)

Related Skills

- Programming Skills: Python, C++, Bash(awk, sed, etc.), LaTex, git, Pytorch, Tensorflow

Awards

- 2016 SJTU Academic Excellence Scholarship Class-A (Top 15%)
- 2017 CS Graduates Education & Development Fund and Yang Yuanqing Education Fund (Top-3 in all graduate students in CS Department).

Academic involvement

- 2016 UAI Travel Award & Volunteer
- 2016 NIPS (Neural Information Processing Systems) Reviewer (5 papers)
- 2017 AAAI sub-reviewer (2 papers)
- 2017 AISTATS Travel Award
- 2018 AAAI Reviewer (1 paper)
- 2019 Frontiers in Genetics (1 paper)
- Frontiers Bioengineering (1 paper)
- 2020 IEEE Journal of Biomedical and Health Informatics (JBHI) Reviewer (1 paper)
- 2020 IEEE Transactions on Cybernetics reviewer (1 paper)
- 2020 ICCCN (The 29th International Conference on Computer Communications and Networks) reviewer (1 paper).
- 2020 PLOS Computational Biology (1 paper)
- 2020 NeurIPS (6 papers)
- 2021 AAAI (3 papers)
- 2021 IJCAI (Senior Program Committee (SPC) members, 2 papers).
- 2021 ICML (5 papers).
- 2021 IEEE Transactions on Neural Networks and Learning Systems (TNNLS) (1 paper).
- 2021 KDD DLG (Deep Learning on Graphs) Workshop (2 papers).
- 2021 NeurIPS (1 paper).
- 2021 Mathematical Biosciences and Engineering (1 paper).
- 2022 ICLR (1 paper).

- **Organizer of NeurIPS 2021 "AI for Science: Mind the Gaps" Workshop** (`https://ai4sciencecommunity.github.io/`).
- 2022 CHIL (Conference on Health, Inference, and Learning) (3 papers).
- 2022 ICML (1 paper).
- **Organizer of ICML 2022 "AI for Science" Workshop.**
- 2022 ICML AI for Science Workshop (18 paper).

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