MENGYING SUN

(517)-220-3770, sunmeng2@msu.edu

EDUCATION

PhD candidate in Computer Science *2016 - Present*
Michigan State University, GPA: 4.0/4.0
M.S. in Statistics *2015*
Michigan State University, GPA: 3.9/4.0

COURSEWORK

Statistical Methods, Design & Theory of Algorithms, Pattern Recognition & Analysis, Artificial Intelligence, Machine Learning, Data Mining, Parallel Computing, Language and Interaction.

WORK EXPERIENCE

Machine Learning Engineer Intern, FAM Ranking Team, Facebook, Inc., CA *May 2021 - Aug 2021*

- Developed ranking models for feed ads format and achieved 10% relative conversion rate uplift over baseline, meanwhile published a runbook in internal tools for ad creative optimization.

Research Assistant, Weill Cornell Medical School, Cornell University, NY *May 2018 - Aug 2018*
Research Staff, Department of Epidemiology & Biostatistics, MSU, MI *May 2015 - Aug 2016*

RESEARCH & PROJECTS

Multi-objective Molecular Generation and Property Optimization. *Fall 2021 - Spring 2022*

- Proposed a method that efficiently generates molecules with desired properties utilizing two-stage Monte Carlo tree search and transformation rules derived from large compound libraries.
- The proposed method achieves better performance than deep learning based methods on various evaluation metrics and is much more computationally efficient.

Contrastive Learning on Molecular Graph Neural Network Pretraining. *Fall 2020 - Spring 2021*

- Proposed a contrastive learning framework with new augmentation which utilizes local and global domain knowledge to improve molecular representation learning without the presence of labels.
- The proposed method achieves superior performance on a variety of molecular tasks such as binding affinity, response in bioassays, toxicity and adverse reactions.

Robust Collaborative Learning on Noisy Labels. *Fall 2019 - Summer 2020*

- Analyzed the mechanism of *disagreement* and *agreement* among multiple networks w.r.t. gradients and label purity during learning process when the training data is presented with label noise.
- Proposed **R**obust **C**ollaborative **L**earning (**RCL**) framework to deal with noisy labels, by adaptively encouraging *disagreement* in the early stage and *agreement* in the later stage to fully boost selection of clean samples for training. The proposed method achieves state-of-art performance and **significantly** outperforms baselines in large noise settings, on both image and bioinformatics data.

Identify Susceptible Locations in EHR via Adversarial Attacks. *Fall 2017 - Spring 2018*

- Built medical predictive modeling via LSTM using Electronic Health Records (EHR).
- Utilized optimization based adversarial attacks to generate candidate adversarial samples for medical records and designed distance metrics to select the best adversarial sample.
- Obtained individual-level and cohort-level susceptible locations for the medical records across each time stamp and measurement using large scale electronic health records MIMIC-III.

SELECTED PUBLICATIONS

1. **Mengying Sun**, Jing Xing, Han Meng, Huijun Wang, Bin Chen, Jiayu Zhou. "MolSearch: Search-based Multi-objective Molecular Generation and Property Optimization." *the 28th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2022**)*, submitted.

2. **Mengying Sun**, Jing Xing, Huijun Wang, Bin Chen, Jiayu Zhou. "MoCL: Contrastive Learning on Molecular Graphs with Multi-level Domain Knowledge." *the 27th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2021**)*.

3. Boyang Liu, **Mengying Sun**, Ding Wang, Pang-Ning Tan, Jiayu Zhou. "Learning Deep Neural Networks under Agnostic Corrupted Supervision." *Proceedings of the 38th International Conference on Machine Learning (**ICML 2021**)*.

4. **Mengying Sun**, Jing Xing, Bin Chen, Jiayu Zhou. "Robust Collaborative Learning with Noisy Labels." *20th IEEE International Conference on Data Mining (**ICDM 2020**)*.

5. **Mengying Sun**, Fei Wang, Olivier Elemento, Jiayu Zhou. "Structure-based Drug-Drug Interaction Detection via Expressive Graph Convolutional Networks and Deep Sets." *The 34th AAAI Conference on Artificial Intelligence Student Abstract (**AAAI 2019**)* .

6. **Mengying Sun**, Sendong Zhao, Coryandar Gilvary, Olivier Elemento, Jiayu Zhou. "Graph Convolutional Networks for Computational Drug Development and Discovery." *Briefings in Bioinformatics*.

7. **Mengying Sun**, Fengyi Tang, Jinfeng Yi, Fei Wang, Jiayu Zhou. "Identify Susceptible Locations in Medical Records via Adversarial Attacks on Deep Predictive Models." *the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2018**)*

8. **Mengying Sun**, Inci M. Baytas, Zhangyang Wang, Jiayu Zhou. "Subspace Network: Deep Multi-Task Censored Regression for Modeling Neurodegenerative Diseases." *the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2018**)*

9. **Mengying Sun**, Gustavo de los Campos. "Locally Dependent Screening: A strategy for developing an accurate genomic predictor using Big Data ." *The ASHG 2018 Annual Meeting*, Poster.

10. Qi Wang, **Mengying Sun**, Liang Zhan, Paul Thompson, Shuiwang Ji, and Jiayu Zhou. "Multi-Modality Disease Modeling via Collective Deep Matrix Factorization." *In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1155-1164. ACM, 2017*.

HONORS & AWARDS

College of Engineering Fellowship: *2017, 2019* **Student Travel award**: *KDD 2018, ICDM 2020*

TEACHING & TUTORIAL EXPERIENCES

FS 2020 **Teaching Assistant**, CSE 404 *Introduction to Machine Learning*, MSU
SS 2018 **Teaching Assistant**, CSE 802 *Pattern Recognition & Analysis*, MSU

PROFESSIONAL SERVICE & OUTREACH ACTIVITIES

Dec 2019 **Invited talk** on *Advances in Deep Learning and its Applications on Drug Discovery*, College of Human Medicine, MSU, Grand Rapids, Michigan
Nov 2019 **Invited talk** on *Statistical Computational Methods in Quantitative Genetics*, University of Pittsburgh, Pennsylvania

TECHNICAL STRENGTHS

Computer Languages	Python, R, Shell, SAS, SQL, LaTeX.
Packages & Tools	Tensorflow, PyTorch, Keras, Scikit-learn, NumPy, SciPy, git, Pandas, Seaborn, Matplotlib, Anaconda, AWS.