Xiaohui Chen

Email : xiaohui.chen@tufts.edu
Phone : 339-545-6204
Github : Xiaohui9607
Linkedin : www.linkedin.com/in/xiaohui-chen

Bio

I am a 2nd-year Ph.D. student at Tufts University advised by Prof. Li-ping Liu. My research interest lies in general machine learning, with emphasis on generative modeling, variational inference and graph learning. Before this, I was interested in domain adaptation and generalization, computer vision and robotics.

Education

Tufts University *Sep. 2021 - present*

Ph.D Candidate in Computer Science

Tufts University *Sep. 2019 - May 2021*

M.S. in Data Science

Professional Experiences

Machine Learning Research Intern Bytedance, Bellevue, USA *Jun. 2022 - Sep. 2022*

Project : data sub-sampling algorithm for recommender system.

- Developed data sub-sampling algorithm to boost CTR model iteration.
- Utilized structural information in user-item network to estimate sampling score for each instance.
- Reduced uncertainty of the estimated sampling scores via label propagation.
- Achieved significant improvement over the AUC metric of the CTR model compared to other sub-sampling algorithms.

Student Research Assistant Tufts University, Medford, USA *May 2020 - Sep. 2021*

Topics : graph learning, abnormal detection, advised by Prof. Li-Ping Liu.

- Developed auto-regressive graph generation models by optimizing the variational lower-bound of the generation orders.
- Improved the AUC scores over anomaly detection tasks by ensembling auto-encoders along with adversarial training.

Topics : video generation in robotics, advised by Prof. Jivko Sinapov and Karen Panetta.

- Developed method that predicts next video frame based on inputs from different modality under robot-object interaction scenario.
- Implemented Partial Derivatives of Boolean Functions(PDBF) algorithm for edge detection using C and C++.

Machine Learning Engineer Cobot Technology, Wuhan, China *Oct. 2018 - Jun. 2019*

Projects : domain adaptation for data reusability, segmentation for anomaly detection.

- Improved classification accuracy by 18% on new domain data by retraining Variational Auto-encoder with unlabeled new domain data and labeled old domain data.
- Developed model based on MobileNet to detect abnormal screws in semi-finished cellphones with 95.3 recall score.

Software Engineer Intern SuperMap Software, Beijing, China *Jun. 2017 - Sep. 2017*

Projects : bicycle sharing system.

- Implemented the query and visualization functions of bicycle location and trajectory.
- Trained an RNN to predict trajectory (79% accuracy), and plan for bike parking using the predictions.

Publications

* Equal contribution

Preprints

1. **Chen, X.**, Li, Y., Zhang, A., Liu, L. NVDiff : Graph Generation through the Diffusion of Node Vectors.

Conferences

1. **Chen, X.**, Wang, Y., Du, Y., Liu, L. On Normalization in Self-supervised Transformers, Conference on Neural Information Processing Systems (NeurIPS) 2023, under review.
2. **Chen, X.**, Sun, J., Wang, T., Guo, R., Liu, L., Zhang, A. Graph-Based Model-Agnostic Data Subsampling for Recommendation Systems, Conference on Knowledge Discovery and Data Mining (SIGKDD) 2023, under review.
3. **Chen, X.**, He, J., Han, X., Liu, L. Efficient and Degree-Guided Graph Generation via Discrete Diffusion Modeling, International Conference on Machine Learning (ICML) 2023
4. **Chen, X.***, Han, X.*, Hu, J., Ruiz, F., and Liu, L. Order Matters : Probabilistic Modeling of Node Sequence for Graph Generation, International Conference on Machine Learning (ICML) 2021.
5. **Chen, X.***, Hosseini, R.*, Panetta, K., and Sinapov, J. A Framework for Multisensory Foresight for Embodied Agents, IEEE International Conference on Robotics and Automation (ICRA) 2021.
6. **Chen, X.***, Han, X.*, and Liu, L. GAN Ensemble for Anomaly Detection, Association for the Advancement of Artificial Intelligence (AAAI) 2021.

Journals

1. Han, X., **Chen, X.**, Ruiz, F., and Liu, L. Fitting Autoregressive Graph Generative Models through Maximum Likelihood Estimation, Journal of Machine Learning Research (JMLR)
2. **Chen, X.***, Chen, X.*, Liu, L. Interpretable Node Representation with Attribute Decoding, Transactions on Machine Learning Research (TMLR)

Professional Services

Conference Reviewer

- Association for the Advancement of Artificial Intelligence (AAAI) *2023*
- International Conference on ArtificialIntelligence and Statistics (AISTATS) *2023*
- Conference on Neural Information Processing Systems (NeurIPS) *2023*

Teaching Experiences

Tufts University

- Introduction to Machine Learning, Graduate Level *Fall 2020*
- Statistical Pattern Recognition, Graduate Level *Spring 2023*

Skills

Programming Languages : Python, C++, C, Shell, Matlab, C#, Java, Javascript.

Libraries & Frameworks : Pytorch, Scikit-learn, OpenCV, ROS.

Tools : Git, Latex, HPC, GDB, Vim, Docker, Valgrind, AWS.

Selected Courses

High Dimension Probability ; Information Theory ; Stochastic Process ; Statistical Pattern Recognition ; Convex Optimization ; Bayesian Deep Learning ; Deep Learning ; Reinforcement Learning ; Natural Language Processing.