Lei Xu

PhD Student at MIT EECS

32 Vassar St, Room 32-D712, Cambridge, MA 02139

+1 (617) 909-2935 | leix@mit.edu | www.citr.me | leix28

Research Interests: Natural Language Processing, Machine Learning, Deep Learning.

Education

Ph.D, Dept of Electronical Engineering and Computer Science(EECS), MIT GPA: 4.0/4.0 — *Cambridge, MA, USA*

Supervised by Dr. Kalyan Veeramachaneni — *Aug. 2022*

Master, Dept of Electronical Engineering and Computer Science(EECS), MIT GPA: 4.0/4.0 — *Cambridge, MA, USA*

Supervised by Dr. Kalyan Veeramachaneni — *Feb. 2020*

Bachelor, Dept of Computer Science and Technology, Tsinghua University GPA: 93/100 — *Beijing, China*

Supervised by Prof. Zhiyuan Liu — *Jul. 2017*

Exchange, Computer Science Department, Carnegie Mellon University — *Pittsburgh, PA, USA*

Supervised by Prof. J. Zico Kolter — *Jun. 2016 - Sept. 2016*

Experience

PhD Research Assistant | DAI Lab, MIT — *Cambridge, MA, USA*

Adversarial Attack and Defense on Text Classifiers (PhD Thesis) — *Nov. 2019 - Jun. 2022*

- Defined a critique score which synthesizes the similarity, fluency and misclassification properties of a sentence, and designed a rewrite and rollback algorithm to generate high-quality adversarial sentences. (Paper submitted to ACL Rolling Review 2022 June.)
- Designed single-word adversarial perturbation attack which achieves comparable attack success rate but much more efficient. Also proposed metrics to quantify classifier robustness under single-word attack. (Paper submitted to COLING 2022)
- Designed an in-situ (test-time) data augmentation method to defend adversarial attacks. (Paper accepted to AdvML Workshop @ KDD 2022.)
- Explored the universal vulnerability of prompt-based classifiers under backdoor and adversarial attack setups. Designed algorithms that can inject or find triggers which cause misclassification on any input text. (Paper accepted to Findings of NAACL 2022.)
- Implemented an open-sourced library–Fibber–to benchmark adversarial attack and defense methods and facilitate future research.

Synthetic Tabular Data Generation use Deep Generative Models (Master Thesis) — *Nov. 2018 - May. 2019*

- Implemented an open-sourced library–SDGym–to thoroughly benchmark existing statistical and neural synthetic data generation models.
- Designed a conditional tabular generative adversarial network (CTGAN) to address the imbalanced distribution of discrete variables and multi-modality of continuous variables. It achieves state-of-the-art performance on SDGym. (Paper accepted to NeurIPS 2019.)

Video Watermarking using Neural Networks — *Sept. 2017 - Jan. 2019*

- Collaborated with a Master student on adding imperceptible watermarks to images and videos using generative adversarial networks.

MLFriend: Automated Problem Discovery — *Nov. 2017 - Nov. 2018*

- Investigated popular event-driven datasets, then formalized a novel data science task - prediction problem discovery.
- Developed an end-to-end pipeline to automatically generate, evaluate, recommend prediction problems on event-driven datasets.

Software Engineer (Hosted by Ji Xue) | Google — *New York, NY, USA*

COVID-19 Recovery Detection — *May 2020 - Aug. 2020*

- Forecasted the recovery time of ads supply and demand using a linear model.
- Implemented the data pipline and showed the forecast on the team's dashboard.

Research Engineer (Hosted by Qingqing Huang) | Google — *Mountain View, CA, USA*

Context-sensitive Matrix Factorization for Recommendation System — *May 2019 - Aug. 2019*

- Designed and implemented a context-sensitive matrix factorization (CFac) method, and achieved significant accuracy improvement.
- Worked closely with my hosts and other folks in the team to apply CFac in the BERT distilling project.

Software Engineer (Hosted by Jiwei Li) | Shannon.AI — *Beijing, China*

Chinese A-Share Stock Question-Answering System — *Jun. 2018 - Aug.2018*

- Developed a module for the main product, a QA system for A-share stocks, with the developing team of 10.

Annual Report Data Mining

- Suggested the roadamp and designed software framework for this research project.
- Implemented a rule-based algorithm to accurately extract production and sales data from companies' annual reports.

Undergrad Research Assistant | THUNLP Lab, Tsinghua University — *Beijing, China*

Nerual Text Summarization — *Nov. 2016 - May. 2017*

- Implemented the sequence-to-sequence summarization model, and trained the model on GPU.
- Open-sourced *THUNLP/TensorFlow-Summarization* repo on GitHub, winning 398 stars so far.
- Improved the model by explicitly incorporating keywords into the loss function, getting significant improvement.

Character Word Embeddings of Chinese *Sept. 2014 - Feb. 2015*

- Designed a joint learning algorithm for Chinese character and word embeddings to improve embedding quality for low-frequency words. (Paper accepted to IJCAI 2015.)

Publications

- **Lei Xu**, Yangyi Chen, Ganqu Cui, Hongcheng Gao, Zhiyuan Liu, *Exploring the Universal Vulnerability of Prompt-based Learning Paradigm*, Findings of NAACL, 2022.
- **Lei Xu**, Laure Berti-Equille, Alfredo Cuesta-Infante, Kalyan Veeramachaneni, *Test-Time Augmentation for Defending Against Adversarial Attacks on Text Classifiers*, AdvML Workshop @ KDD, 2022.
- **Lei Xu**, Kalyan Veeramachaneni, *Attacking Text Classifiers via Sentence Rewriting Sampler*, Preprint, 2020.
- **Lei Xu**, Maria Skoularidou, Alfredo Cuesta-Infante, Kalyan Veeramachaneni, *Modeling Tabular Data using Conditional GAN*, NeurIPS, 2019. (Citation: 263)
- **Lei Xu**, Kalyan Veeramachaneni, *Synthesizing tabular data using generative adversarial networks*, Preprint, 2018. (Citation: 107)
- Brandon Amos, **Lei Xu**, J. Zico Kolter, *Input Convex Neural Networks*, ICML, 2017. (Citation: 275)
- Xinxiong Chen*, **Lei Xu***, Zhiyuan Liu, Maosong Sun, Huanbo Luan. *Joint Learning of Character and Word Embeddings*, IJCAI, 2015. (* equal contribution; Citation: 332)
- Kevin Alex Zhang, Alfredo Cuesta-Infante, **Lei Xu**, Kalyan Veeramachaneni, *SteganoGAN: High capacity image steganography with GANs*, Preprint, 2019. (Citation: 72).
- Shubhra Kanti Karmaker Santu, Md. Mahadi Hassan, Micah J. Smith, **Lei Xu**, ChengXiang Zhai, Kalyan Veeramachaneni, *AutoML to Date and Beyond: Challenges and Opportunities*, ACM Computing Surveys, 2021.

Teaching Experiences

2021 **Teaching Assistant** 6.86X Machine Learning with Python (MIT)

2017 **Teaching Assistant** Media Programing (Tsinghua University)

Awards/Skills

Awards 2017 Outstanding Graduates of Beijing

Programming Languages C/C++, Java, Python, Tensorflow, PyTorch, Shell, HTML, Javascript, SQL.

English (fluent), Chinese (native).