

Xiaodi Li

401 Broadway Avenue North, Rochester, MN 55906
(469) 380-4852 | xiaodi.li@utdallas.edu | <https://xiaodili.org/>

Education

Department of Electrical and Computer Engineering, the University of Texas at Dallas	Richardson, TX
Ph.D. of Computer Engineering	Aug. 2018 - Aug. 2024
Dissertation: <i>Advanced Approaches in NLP and Security: Addressing Catastrophic Forgetting through Continual Learning and Resolving Data Imbalance in Semi-Supervised Settings</i>	
College of Mechanical Engineering, Donghua University	Shanghai, P.R.China
Bachelor of Mechanical Engineering	Jun. 2018
Graduation Thesis: <i>Detection and Extraction of Texture Defect Features on Leather Surface</i>	
Nominate for the Honor of Excellent Undergraduate Graduation Thesis	

Research Interests

- ◆ My current research interests are **Artificial Intelligence, Machine Learning, and Natural Language Processing.**

Selected Publications

- ◆ Xiaodi Li, Dingcheng Li, Rujun Gao, Mahmoud Zamani, and Latifur Khan. "[LSEBMCL: A Latent Space Energy-Based Model for Continual Learning.](#)" *In the 7th International Conference on Artificial Intelligence in Information and Communication (ICAIC 2025).*
- ◆ Xiaodi Li, Niamat Zawad, Patrick T. Brandt, Javier Osorioc, Vito D'Orazio, and Latifur Khan. "[ConfliLPC: Logits and Parameter Calibration for Political Conflict Analysis in Continual Learning.](#)" *In 2024 IEEE International Conference on Big Data (BigData), pp. 6320-6329. IEEE, 2024.*
- ◆ Saquib Irtiza, Xiaodi Li, Mahmoud Zamani, Latifur Khan, and Kevin W. Hamlen. "[VulPrompt: Prompt-Based Vulnerability Detection Using Few-Shot Graph Learning.](#)" *In IFIP Annual Conference on Data and Applications Security and Privacy, DBSec 2024, pp. 221-240. Cham: Springer Nature Switzerland, 2024.*
- ◆ Xiaodi Li, Latifur Khan, Mahmoud Zamani, Shamila Wickramasuriya, Kevin Hamlen, and Bhavani Thuraisingham. "[Con2Mix: A semi-supervised method for imbalanced tabular security data.](#)" *Journal of Computer Security, vol. 31, no. 6, pp. 705-726, 2023.*
- ◆ Xiaodi Li, Md Delwar Hossain, Hideya Ochiai, and Latifur Khan. "[2MiCo: A Contrastive Semi-Supervised Method with Double Mixup for Smart Meter Modbus RS-485 Communication Security.](#)" *In 2023 IEEE 9th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS), pp. 30-39. IEEE, 2023.*
- ◆ Xiaodi Li, Zhuoyi Wang, Dingcheng Li, Latifur Khan, and Bhavani Thuraisingham. "[LPC: A Logits and Parameter Calibration Framework for Continual Learning.](#)" *In Findings of the Association for Computational Linguistics: EMNLP 2022, pp. 7142-7155. 2022.*
- ◆ Xiaodi Li, Latifur Khan, Mahmoud Zamani, Shamila Wickramasuriya, Kevin W. Hamlen, and Bhavani Thuraisingham. "[MCoM: A Semi-Supervised Method for Imbalanced Tabular Security Data.](#)" *In Data and Applications Security and Privacy XXXVI: 36th Annual IFIP WG 11.3 Conference, DBSec 2022, Newark, NJ, USA, July 18-20, 2022, Proceedings, pp. 48-67. Cham: Springer International Publishing, 2022.*
- ◆ Simin Chen, Soroush Bateni, Sampath Grandhi, Xiaodi Li, Cong Liu, and Wei Yang. "[DENAS: Automated](#)

[Rule Generation by Knowledge Extraction from Neural Networks.](#)" *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE)*, 2020. doi:10.1145/3368089.3409733.

- ◆ Yueming Wu, Xiaodi Li, Deqing Zou, Wei Yang, Xin Zhang, and Hai Jin. "[MalScan: Fast Market-Wide Mobile Malware Scanning by Social-Network Centrality Analysis.](#)" *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2019. doi:10.1109/ase.2019.00023.

Research Experience

Big Data Analytics and Management Lab, The University of Texas at Dallas	Richardson, TX
‘A New Semi-Supervised Learning Methodology on Tabular Security Data Sets’	<i>Mar. 2022 - Apr. 2022</i>
◆ Propose a novel triplet mixup data augmentation method that can reduce the data imbalance problem in large, sparsely labeled data sets	
◆ The first semi-supervised learning framework on tabular security data sets	
◆ Achieve state-of-the-art performance on all the experiments, which demonstrate the superiority of our proposed method for cybersecurity data domains	
‘A Logits and Parameter Calibration Framework on Continual Learning to Reduce Catastrophic Forgetting’	<i>Dec. 2020 - Jun. 2022</i>
◆ Propose a new continual learning framework LPC, which can reduce catastrophic forgetting effectively	
◆ Develop a new mechanism by calibrating the logits and parameters with target drifting from previous tasks to current tasks, thereby alleviating the catastrophic forgetting during the model updating	
◆ Combined with the existing regularization-based approach, our model achieves state-of-the-art performance while addressing the old knowledge forgetting without data storage	
Dallas Intelligent Software and Security Lab, The University of Texas at Dallas	Richardson, TX
‘Perform Android Malware Classification with a Lightweight Graph-based Approach’	<i>Jan. 2019 - Feb.2019</i>
◆ Extract API and manifest features from downloaded APKs	
◆ Test the efficiency of the extracted features using SVM method	
Key Lab of Cloud Computing and Intelligent Technology, Southwest Jiaotong University	Chengdu, P.R.China
‘Detecting Abnormalities with DP (density peaks) Clustering Method’	<i>Sep. 2017 - Feb.2018</i>
◆ Design a distribution method based on DP Algorithm to find abnormal points	
◆ Test our method with MATLAB using large scale of sparse dataset	
WSN Lab, Sichuan University	Chengdu, P.R.China
‘Mapping Client Messages to a Unified Data Model with MfeCNN’	<i>Dec. 2016 - Aug. 2017</i>
◆ Employ a tool named cTAKES (clinical Text Analysis Knowledge Extraction System) from Apache to extract multi-modal features from the input sentences	
◆ Create word embeddings for input features	
‘Improving Energy Efficiency of Neighbor Discovery in WSNs with ODM’	<i>2016 Summer</i>
◆ Do several derivations of formulas such as the general state expression, $M(c)$ and the number of Probe slots	
◆ Simulate our method in the dynamic case	

Work Experience

Mayo Clinic	Rochester, MN
Postdoctoral Research Fellow – Large Language Model (LLM) and Artificial Intelligence	<i>Sep. 2024 - Present</i>
◆ Conduct research in the Department of Artificial Intelligence and Informatics	
The University of Texas at Dallas	Richardson, TX
Research Assistant	<i>Aug. 2021 - Aug. 2024</i>
◆ Assist research in the Big Data Analytics and Management Lab	

Teaching Assistant *Aug. 2018 - Aug. 2021*

◆ Assist teaching for the following courses: Database System, Semantic Web, Introduction to Machine Learning, Data Structures and Introduction to Algorithmic Analysis, Computer Science II, and Machine Learning in Cyber Security

Tektronix Company **Shanghai, P.R.China**

Software Test Engineer Intern *Jul. 15th, 2015 - Aug. 25th, 2015*

◆ Test the UI for the main software system of the company; record errors occurred during the test; write test reports and send them to software design engineers

Shanghai Aircraft Manufacturing Factory, Roewe Group Shanghai Base Port, and China International Industrial Expo **Shanghai, P.R.China**

Cognition Practice *2014 - 2015*

Leadership Experience

The Future of Illusion - Robot Community of Donghua University **Shanghai, P.R.China**

Vice-Chairman *Mar. 2015 - Nov. 2015*

The Future of Illusion - Robot Community of Donghua University **Shanghai, P.R.China**

Propaganda Department Minister *Sep. 2014 - Mar. 2015*

Honors & Awards

The Excellence in Education Foundation Doctoral Fellowship, the University of Texas at Dallas **Richardson, TX**

2023-2024 academic year *Apr. 10th, 2023*

The Louis Beecherl, Jr. Graduate Fellowship, the University of Texas at Dallas **Richardson, TX**

2021-2022 academic year *Jul. 1st, 2021*

Volunteer Work

ACM CODASPY 2019 **Dallas, TX**

Volunteer *Mar. 25th, 2019 - Mar. 27th, 2019*

The Third Session of the Original Music Contest of Shanghai College Students **Shanghai, P.R.China**

Volunteer *Oct. 21st, 2014*

The First Session of Chinese Robot Tourism Competition **Shanghai, P.R.China**

Volunteer *May. 12th, 2014*

Skills/Interests

Programming Languages: Python, JAVA, C/C#, SQL, HTML, XML, JavaScript, CSS, Assembly Language, and Matlab

Languages: Chinese (native) and English (proficient)

Interests: Thinking, Reading, and Writing; Fitness; and Playing Guitar