

Dongkuan (DK) Xu

CONTACT INFORMATION	College of Information Sciences and Technology The Pennsylvania State University <i>E-mail:</i> dux19@psu.edu <i>Mobile:</i> +1-814-699-0860	<i>Web:</i> https://personal.psu.edu/dux19/ <i>Google Scholar:</i> Link <i>Twitter:</i> https://twitter.com/DongkuanXu <i>Updated:</i> Jan 12th, 2022
RESEARCH INTERESTS	I am interested in efficient AI, including parameter efficiency, data efficiency, and computation efficiency. My research investigates how to improve the efficiency of deep learning systems to achieve a Pareto optimality between resources (e.g., parameter, data, computation) and performance (e.g., inference, training). My long-term research goal is to free AI from the parameter-data-computation hungry beast and to democratize AI to serve a broader area and population. <ul style="list-style-type: none">• Parameter Efficiency: Neural Architecture Search, Pruning, Knowledge Distillation• Data Efficiency: Few-shot Learning, Self-supervised Learning• Computation Efficiency: Weight-sharing Learning, Reduced-cost Training• Domains: NLP, Computer Vision, Social Science, Neuroinformatics, Finance• Applications: Model Compression for Memory-limited Devices, Privacy-preserving Deep Learning, Energy-efficient AI Systems, Efficient AI for Science (Biology/Chemistry/Physics)	
EDUCATION	The Pennsylvania State University , PA, USA <ul style="list-style-type: none">• Ph.D. in Informatics• College of IST Award for Excellence in Teaching Support (top 2)• Adviser: Xiang Zhang University of Chinese Academy of Sciences , Beijing, China <ul style="list-style-type: none">• M.S. in Optimization• Chinese Academy of Sciences President's Fellowship (the most prestigious award)• Adviser: Yingjie Tian Renmin University of China , Beijing, China <ul style="list-style-type: none">• B.E. in Information Management and System	Aug 2017 - Present Sep 2014 - May 2017 Sep 2010 - May 2014
PROFESSIONAL EXPERIENCE (INDUSTRY)	Microsoft Research (MSR) , Redmond, WA <ul style="list-style-type: none">• Research Intern, Mentors: Subho Mukherjee, Xiaodong Liu, Debadepta Dey, Ahmed H. Awadallah, Jianfeng Gao• Project: Task-agnostic Auto-Transformer Search Moffett.AI , Los Altos, CA <ul style="list-style-type: none">• Research Intern, Mentor: Ian En-Hsu Yen, Co-founder• Project: Knowledge-aware Pruning of Pre-trained Models• Project: Data-free Model Compression• Publication: NAACL 2021 and a U.S. patent NEC Labs America , Princeton, NJ <ul style="list-style-type: none">• Research Intern, Mentor: Wei Cheng• Project: Knowledge Transfer in Multi-Task Learning• Project: Trend Learning in Multivariate Time Series• Publication: AAAI 2021, AAAI 2020 NEC Labs America , Princeton, NJ <ul style="list-style-type: none">• Research Intern, Mentor: Wei Cheng, Senior Researcher• Project: Contrastive Anomaly Detection• Publication: SDM 2021	May 2021 - Aug 2021 May 2020 - Jul 2020 May 2019 - Jan 2020 May 2018 - Aug 2018

PROFESSIONAL EXPERIENCE (ACADEMIA)	<p>The Pennsylvania State University Aug 2017 - Present</p> <ul style="list-style-type: none"> • Graduate Research Assistant, Adviser: Xiang Zhang • Thesis: Efficient Deep Learning Methods for Temporal Modeling • Publication: NeurIPS'21, 2×AAAI'21, SDM'21, AAAI'20, IJCAI'19, ICDM'19, SDM'19 <p>Chinese Academy of Sciences, Beijing, China Sep 2014 - May 2017</p> <ul style="list-style-type: none"> • Graduate Research Assistant, Adviser: Yingjie Tian • Thesis: Efficient Multi-instance Learning • Publication: PR'18, IJCNN'17, ITQM'16, AODS'15 <p>Renmin University of China, Beijing, China Sep 2012 - May 2014</p> <ul style="list-style-type: none"> • Undergraduate Research Assistant, Adviser: Wei Xu • Thesis: Ensemble Forecasting Model for Time Series Data • Publication: IJGEI'15, HICSS'14
HONORS AND AWARDS	<p>Doctor of Philosophy (Ph.D.)</p> <ul style="list-style-type: none"> • College of IST Award for Excellence in Teaching Support (top 2) 2018 • IST Travel Award (Fall) 2021 • NAACL Scholarship 2021 • SIAM Student Travel Award 2021 • IST Travel Award (Spring) 2021 • KDD Student Registration Award 2020 • AAAI Student Scholarship 2020 • IST Travel Award (Fall) 2020 • IST Travel Award (Spring) 2019 <p>Master of Science (M.S.)</p> <ul style="list-style-type: none"> • Chinese Academy of Sciences President's Fellowship (top 1) 2016 • Information Technology and Quantitative Management (ITQM) Best Paper 2016 • National Graduate Scholarship, China (2% in university) 2016 • Graduate Student Academic Scholarship 2017 • Graduate Student Academic Scholarship 2016 • Graduate Student Academic Scholarship 2015 <p>Bachelor of Engineering (B.E.)</p> <ul style="list-style-type: none"> • First-class Scholarship of Sashixuan Elite Fund, China (5% in university) 2014 • Kwang-hua Scholarship of RUC, China 2014 • Second-class Scholarship of Excellent Student Cadre 2014 • Meritorious Winner in Mathematical Contest in Modeling, USA 2013 • First-class Scholarship of Social Work and Volunteer Service of RUC 2013
PUBLICATION SUMMARY	<p>Published: 26 papers, 13 first-authored papers, 8 preprints, and 8 filed patents. Impact: 1200+ citations, h-index: 9 (as of Jan 12th 2022)</p>
PEER-REVIEWED CONFERENCE AND JOURNAL PAPERS	<p>[1] Dongkuan Xu, Wei Cheng, Dongsheng Luo, Haifeng Chen, and Xiang Zhang. InfoGCL: Information-Aware Graph Contrastive Learning [C]. The 35th Conference on Neural Information Processing Systems (NeurIPS 2021) Acceptance rate: $2372/9122=26.0\%$</p> <p>[2] Dongkuan Xu, Ian En-Hsu Yen, Jinxi Zhao, and Zhibin Xiao. Rethinking Network Pruning - under the Pre-train and Fine-tune Paradigm [C]. 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2021) Acceptance rate: $477/1797=26.5\%$</p>

- [3] Xin Dong, Yaxin Zhu, Zuohui Fu, **Dongkuan Xu**, and Gerard de Melo. Data Augmentation with Adversarial Training for Cross-Lingual NLI [C]. The 59th Annual Meeting of the Association for Computational Linguistics (**ACL 2021**)
Acceptance rate: $714/3350=21.3\%$
- [4] **Dongkuan Xu**, Wei Cheng, Jingchao Ni, Dongsheng Luo, Masanao Natsumeda, Dongjin Song, Bo Zong, Haifeng Chen, and Xiang Zhang. Deep Multi-Instance Contrastive Learning with Dual Attention for Anomaly Precursor Detection [C]. The 21th SIAM International Conference on Data Mining (**SDM 2021**)
Acceptance rate: $85/400=21.3\%$
- [5] **Dongkuan Xu**, Wei Cheng, Xin Dong, Bo Zong, Wenchao Yu, Jingchao Ni, Dongjin Song, Xuchao Zhang, Haifeng Cheng, and Xiang Zhang. Multi-Task Recurrent Modular Networks [C]. The 35th AAAI International Conference on Artificial Intelligence (**AAAI 2021**)
Acceptance rate: $1692/7911=21.4\%$
- [6] **Dongkuan Xu**, Junjie Liang, Wei Cheng, Hua Wei, Haifeng Cheng, and Xiang Zhang. Transformer Style Relational Reasoning with Dynamic Memory Updating for Temporal Network Modeling [C]. The 35th AAAI International Conference on Artificial Intelligence (**AAAI 2021**)
Acceptance rate: $1692/7911=21.4\%$
- [7] Hua Wei, **Dongkuan Xu**, Junjie Liang, and Zhenhui Li. How Do We Move: Modeling Human Movement with System Dynamics [C]. The 35th AAAI International Conference on Artificial Intelligence (**AAAI 2021**)
Acceptance rate: $1692/7911=21.4\%$
- [8] Junjie Liang, Yanting Wu, **Dongkuan Xu**, and Vasant Honavar. Longitudinal Deep Kernel Gaussian Process Regression [C]. The 35th AAAI International Conference on Artificial Intelligence (**AAAI 2021**)
Acceptance rate: $1692/7911=21.4\%$
- [9] Dongsheng Luo, Wei Cheng, **Dongkuan Xu**, Wenchao Yu, Bo Zong, Haifeng Chen, and Xiang Zhang. Parameterized Explainer for Graph Neural Network [C]. The 34th Conference on Neural Information Processing Systems (**NeurIPS 2020**)
Acceptance rate: $1900/9454=20.1\%$
- [10] Xin Dong, Yaxin Zhu, Yupeng Zhang, Zuohui Fu, **Dongkuan Xu**, Sen Yang, and Gerard de Melo. Leveraging Adversarial Training in Self-Learning for Cross-Lingual Text Classification [C]. The 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (**SIGIR 2020**)
Acceptance rate: $300/1062=28.2\%$
- [11] **Dongkuan Xu**, Wei Cheng, Bo Zong, Dongjin Song, Jingchao Ni, Wenchao Yu, Yanchi Liu, Haifeng Chen, and Xiang Zhang. Tensorized LSTM with Adaptive Shared Memory for Learning Trends in Multivariate Time Series [C]. The 34th AAAI International Conference on Artificial Intelligence (**AAAI 2020**)
Acceptance rate: $1591/7737=20.6\%$
- [12] Junjie Liang, **Dongkuan Xu**, Yiwei Sun, and Vasant Honavar. Longitudinal Multi-Level Factorization Machines [C]. The 34th AAAI International Conference on Artificial Intelligence (**AAAI 2020**)
Acceptance rate: $1591/7737=20.6\%$
- [13] **Dongkuan Xu**, Wei Cheng, Dongsheng Luo, Yameng Gu, Xiao Liu, Jingchao Ni, Bo Zong, Haifeng Chen, and Xiang Zhang. Adaptive Neural Network for Node Classification in Dynamic Networks [C]. The 19th IEEE International Conference on Data

Mining (**ICDM 2019**)

Acceptance rate: 183/930=19.7%

- [14] **Dongkuan Xu**, Wei Cheng, Dongsheng Luo, Xiao Liu, and Xiang Zhang. Spatio-Temporal Attentive RNN for Node Classification in Temporal Attributed Graphs [C]. The 28th International Joint Conference on Artificial Intelligence (**IJCAI 2019**)
Acceptance rate: 850/4752=17.9%
- [15] **Dongkuan Xu**, Wei Cheng, Bo Zong, Jingchao Ni, Dongjin Song, Wenchao Yu, Yuncong Chen, Haifeng Chen, and Xiang Zhang. Deep Co-Clustering [C]. The 19th SIAM International Conference on Data Mining (**SDM 2019**)
Acceptance rate: 90/397=22.7%
- [16] Jingchao Ni, Shiyu Chang, Xiao Liu, Wei Cheng, Haifeng Chen, **Dongkuan Xu**, and Xiang Zhang. Co-Regularized Deep Multi-Network Embedding [C]. The 27th International Conference on World Wide Web (**WWW 2018**)
Acceptance rate: 170/1175=14.5%
- [17] **Dongkuan Xu**, Wei Zhang, Jia Wu, and Yingjie Tian. Multiple Instance Learning Based on Positive Instance Graph [J]. arXiv preprint arXiv: 1612.03550 (2016)
- [18] Yingjie Tian, **Dongkuan Xu**, and Chunhua Zhang. A Review of Multi-Instance Learning Research [J]. Operations Research Transactions, 2018, 02: 1-17
- [19] **Dongkuan Xu**, Jia Wu, Dewei Li, Yingjie Tian, Xinquan Zhu, and Xindong Wu. SALE: Self-Adaptive LSH Encoding for Multi-Instance Learning [J]. **Pattern Recognition**, 2017 (**7.74 impact factor**)
- [20] Dewei Li, **Dongkuan Xu**, Jingjing Tang, and Yingjie Tian. Metric Learning for Multi-Instance Classification with Collapsed Bags [C]. The 30th IEEE International Joint Conference on Neural Networks (**IEEE IJCNN 2017**)
- [21] **Dongkuan Xu**, Jia Wu, Wei Zhang, and Yingjie Tian. PIGMIL: Positive Instance Detection via Graph Updating for Multiple Instance Learning [C]. arXiv preprint arXiv:1612.03550, 2016
- [22] Dewei Li, Wei Zhang, **Dongkuan Xu**, and Yingjie Tian. Multi-Metrics Classification Machine [C]. International Conference on Information Technology and Quantitative Management (**ITQM 2016**) (**Best Paper Award**)
- [23] **Dongkuan Xu**, and Yingjie Tian. A Comprehensive Survey of Clustering Algorithms [J]. Annals of Data Science, 2015, 2(2): 165-193
- [24] **Dongkuan Xu**, Tianjia Chen, and Wei Xu. A Support Vector Machine-Based Ensemble Prediction for Crude Oil Price with VECM and STEPMS [J]. International Journal of Global Energy Issues, 2015
- [25] **Dongkuan Xu**, Yi Zhang, Cheng Cheng, Wei Xu, and Likuan Zhang. A Neural Network-Based Ensemble Prediction Using PMRS and ECM [C]. The 47th IEEE Hawaii International Conference on System Sciences (**HICSS 2014**)
- [26] **Dongkuan Xu**, Subhabrata Mukherjee, Xiaodong Liu, Debadeepta Dey, Ahmed H. Awadallah, Jianfeng Gao, and Xiang Zhang. Few-shot Task-agnostic Neural Architecture Search for Distilling Large Language Models [C].
- [27] Shaoyi Huang*, **Dongkuan Xu***, Ian En-Hsu Yen, Sung-En Chang, Bingbing Li, Shiyang Chen, Mimi Xie, Hang Liu, and Caiwen Ding. Sparse Progressive Distillation: Resolving Overfitting under Pretrain-and-Finetune Paradigm [C]. arXiv preprint arXiv:2110.08190, 2021

PREPRINTS AND
SUBMISSIONS
(* EQUAL
CONTRIBUTION)

- [28] **Dongkuan Xu**, and Ian En-Hsu Yen. Neural Network Pruning without Retraining [C].
- [29] Junjie Liang, Hua Wei, **Dongkuan Xu**, and Vasant G Honavar. Inducing Clusters Deep Kernel Gaussian Process for Longitudinal Data [C].
- [30] Dongsheng Luo, Wei Cheng, Yingheng Wang, **Dongkuan Xu**, Jingchao Ni, Wenchao Yu, Xuchao Zhang, Yanchi Liu, Haifeng Chen, and Xiang Zhang. Information-Aware Time Series Meta-Contrastive Learning [C].
- [31] Dongsheng Luo, Wei Cheng, **Dongkuan Xu**, Feng Han, Xiao Liu, Wenchao Yu, Haifeng Chen, and Xiang Zhang. Inductive and Efficient Explanations for General Graph Neural Networks [J].

TEACHING
EXPERIENCE

Teaching Assistant at Penn State University

- SRA 268, Visual Analytics Fall 2021
Instructor: Prof. [Mahir Akgun](#)
Course Materials: [Visual Analytics with Tableau](#)
(**Responsible for teaching lab classes of 46 students**)
- SRA 450, Cybercrime and Cyberwar Fall 2021
Instructor: Prof. [John Hodgson](#)
Course Materials: [Cybersecurity: What Everyone Needs to Know](#)
- DS/CMPSC 410, Programming Models for Big Data Spring 2021
Instructor: Prof. [John Yen](#)
Course Materials: [Learning Spark](#)
- SRA 365, Statistics for Security and Risk Analysis Fall 2020
Instructor: Dr. [James Farrugia](#)
Course Materials: [Discovering Statistics Using R](#)
- DS 402, Introduction to Social Media Mining Spring 2020
Instructor: Prof. [Suhang Wang](#)
Course Materials: [Social Media Mining: An Introduction](#)
- SRA 365, Statistics for Security and Risk Analysis Spring 2019
Instructor: Dr. [Katherine Hamilton](#)
Course Materials: [Foundations and Practice of Intermediate Statistics](#)
- IST 210, Organization of Data Fall 2018
Instructor: Prof. [Xiang Zhang](#)
Course Materials: [Database Systems Concepts](#)
(**The Award for Excellence in Teaching Support**)

Guest Lecturer

- COSI 133A, Graph Mining Fall 2021
Brandeis University
Slides: [Link](#)
- COSI 165B, Deep Learning Spring 2021
Brandeis University
Slides: [Link](#)

MENTORING
EXPERIENCE

- [Shaoyi Huang](#), Ph.D. student at University of Connecticut 2021-Present
Research Topic I: Sparse Neural Architecture Search
Research Topic II: Few-shot BERT Distillation

- Tianxiang Zhao, Ph.D. student at Penn State University 2021-Present
Research Topic: Graph Transfer Learning
- Zhenglun Kong, Ph.D. student at Northeastern University 2021-Present
Research Topic: Efficient Auto Vision Transformer Search
- Shanglin Zhou, Ph.D. student at University of Connecticut 2021-Present
Research Topic: Data-free Model Compression
- Bowen Lei, Ph.D. student at Texas A&M University 2021-Present
Research Topic: Robust Sparse Neural Network Training
- Wei Zhang, Undergraduate student at Renmin University of China 2016-2017
Research Topic: Cost-Sensitive Multi-Instance Learning
(Current Position: Ph.D. student at City University of Hong Kong)

PROFESSIONAL
SERVICE

Academic Committee Member

- Machine Learning & Natural Language Processing Community (MLNLP)

Senior Program Committee

- International Joint Conferences on Artificial Intelligence (IJCAI) 2021

Program Committee

- Neural Information Processing Systems (NeurIPS) 2020, 2021
- International Conference on Learning Representations (ICLR) 2021, 2022
- International Conference on Machine Learning (ICML) 2021, 2022
- AAI Conference on Artificial Intelligence (AAAI) 2020, 2021, 2022
- SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2020, 2021, 2022
- Association for Computational Linguistics (ACL) Rolling Review 2022
- North American Chapter of the Association for Computational Linguistics (NAACL) 2021
- Conference on Empirical Methods in Natural Language Processing (EMNLP) 2020, 2021
- International Joint Conferences on Artificial Intelligence (IJCAI) 2020, 2022
- ACM International Conference on Web Search and Data Mining (WSDM) 2022
- SIAM International Conference on Data Mining (SDM) 2022
- European Chapter of the Association for Computational Linguistics (EACL) 2021
- Conference on Information and Knowledge Management (CIKM) 2020, 2021
- Asia-Pacific Chapter of the Association for Computational Linguistics & International Joint Conference on Natural Language Processing (AACL-IJCNLP) 2020
- International Joint Conference on Neural Networks (IJCNN) 2018, 2019, 2020, 2021

Journal Reviewer

- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- IEEE Transactions on Cybernetics
- Information Fusion
- ACM Transactions on Knowledge Discovery from Data (TKDD)
- Pattern Recognition
- Neural Networks
- ACM Transactions on Asian and Low-Resource Language Information Processing
- IEEE Access
- Neural Computation
- Complexity
- Soft Computing
- Journal of Sports Engineering and Technology
- Complex & Intelligent Systems
- Multimedia Tools and Applications
- Big Data

External Conference Reviewer

- AAAI'18-20, ACM CIKM'18-19, Big Data'18, ICDM'18-19, IJCNN'16-17, ITQM'16-17, KDD'18-21, SDM'18-22, TheWebConf (WWW)'20-22, WSDM'20-21

Conference Volunteer

- The Annual Conference of NAACL-HLT, 2021
- Backuping SDM Session Chairs, 2021
- The 35th AAAI Conference on Artificial Intelligence, 2021
- The 26th SIGKDD Conference on Knowledge Discovery and Data Mining, 2020

PATENT & APPLICATIONS

- System and Method for Knowledge-Preserving Neural Network Pruning
Enxu Yan, **Dongkuan Xu**, and Zhibin Xiao
U.S. Patent. 11,200,497. Dec. 2021
- Neural Network Pruning Method and System via Layerwise Analysis
Enxu Yan, **Dongkuan Xu**, and Jiachao Liu
U.S. Patent App. 17/107,046. Nov. 2020
- Bank-Balanced Sparse Activation Feature Maps for Neural Network Models
Enxu Yan, **Dongkuan Xu**, and Jiachao Liu
U.S. Patent App. Sep. 2020
- Unsupervised Multivariate Time Series Trend Detection for Group Behavior Analysis
Wei Cheng, Haifeng Chen, Jingchao Ni, **Dongkuan Xu**, and Wenchao Yu
U.S. Patent App. 16/987,734. Mar. 2021
- Tensorized LSTM with Adaptive Shared Memory for Learning Trends
Wei Cheng, Haifeng Chen, Jingchao Ni, **Dongkuan Xu**, and Wenchao Yu
U.S. Patent App. 16/987,789. Mar. 2021
- Adaptive Neural Networks for Node Classification in Dynamic Networks
Wei Cheng, Haifeng Chen, Wenchao Yu, and **Dongkuan Xu**
U.S. Patent App. 16/872,546. Nov. 2020
- Spatio Temporal Gated Recurrent Unit
Wei Cheng, Haifeng Chen, and **Dongkuan Xu**
U.S. Patent App. 16/787,820. Aug. 2020
- Automated Anomaly Precursor Detection
Wei Cheng, **Dongkuan Xu**, Haifeng Chen, and Masanao Natsumeda
U.S. Patent App. 16/520,632. Feb. 2020

PROFESSIONAL TALKS

- Parameter Efficiency: Democratizing AI at Scale (Slides)
Brandeis University, Waltham, USA, Dec 2021
- Chasing Efficiency of Pre-trained Language Models
Microsoft Research Lab, Redmond, Washington, USA, Jun 2021
- BERT Pruning: Structural vs. Sparse (Slides)
Brandeis University, Waltham, USA, Apr 2021
- BERT, Compression, and Applications (Slides)
Xpeng Motors, Mountain View, USA, Apr 2021
- BERT Architecture and Computation Analysis
Moffett.AI, Los Altos, USA, May 2020.

- Learning Trends in Multivariate Time Series (Slides)
In AAAI, New York, USA, Feb 2020
- Node Classification in Dynamic Networks (Slides)
In ICDM, Beijing, China, Nov 2019
- Anomaly Precursor Detection via Multi-Instance Contrastive Learning
NEC Laboratories America, Princeton, USA, May 2019
- Deep Co-Clustering (Slides)
In SDM, Calgary, Canada, May 2019
- Efficient Multiple Instance Learning (Slides)
NEC Laboratories America, Princeton, USA, May 2018

EXTRACURRICULAR ACTIVITIES	• ACM Student Membership	2021-Present
	• ACL Membership	2021-Present
	• AAAI Student Membership	2019-2021
	• SIAM CAS Student Member	2016-2021
	• Volunteer of Beijing Volunteer Service Federation	2012-2014
	• President of Youth Volunteers Association of School of Information	2012-2013
	• Leader of National Undergraduate Training Programs	2011-2012

SKILLS	• Programming: Python, C, R, MATLAB, SQL, L ^A T _E X
	• Tools: PyTorch, TensorFlow, Keras, Scikit-learn, SPSS

REFEREES **Available upon request**