Chengyu Dong

cdong@eng.ucsd.edu
www.chengyu-dong.me

Research Interest

Efficient and Robust Machine Learning

Education

Ph.D. in Computer Science   Sept. 2021 - Present
Jacobs School of Engineering, University of California, San Diego

M.S. in Computer Science   Sept. 2018 - June 2020
Jacobs School of Engineering, University of California, San Diego

M.S. in Astrometry and Celestial Mechanics   Sept. 2016 - Aug. 2018
School of Astronomy & Space Science, Nanjing University

B.S. in Astronomy   Sept. 2012 - June 2016
Kuang Yaming Honors School (Gifted Youth Program), Nanjing University

Selected Publication

Chengyu Dong, Liyuan Liu and Jingbo Shang. SoTeacher: A Student-oriented Teacher Network Training Framework for Knowledge Distillation. *arXiv:2206.06661*

Chengyu Dong, Liyuan Liu and Jingbo Shang. Label Noise in Adversarial Training: A Novel Perspective to Study Robust Overfitting. *NeurIPS 2022 (Oral presentation)*.

Chengyu Dong, Liyuan Liu and Jingbo Shang. Data Quality Matters For Adversarial Training: An Empirical Study. *arXiv:2102.07437*.

Chengyu Dong, Liyuan Liu, Zichao Li and Jingbo Shang. Towards Adaptive Residual Network Training: A Neural-ODE Perspective. *ICML 2020*.

Dheeraj Mekala, **Chengyu Dong** and Jingbo Shang. LOPS: Learning Order Inspired Pseudo-Label Selection for Weakly Supervised Text Classification. *Findings of EMNLP 2022*.

Zichao Li, Dheeraj Mekala, **Chengyu Dong** and Jingbo Shang. BFClass: A Backdoor-free Text Classification Framework. *Findings of EMNLP 2021*.

Zichao Li, Liyuan Liu, **Chengyu Dong**, and Jingbo Shang. Overfitting or Underfitting? Understand Robustness Drop in Adversarial Training. *arXiv:2010.08034*.

Zihan Wang, **Chengyu Dong** and Jingbo Shang. Average Approximates First Principal Component? An Empirical Analysis on Representations from Neural Language Models. *EMNLP 2021*.

Experience

Research Intern @ Deep Learning Group, Microsoft Research   June 2022 - Present
Supervisor: Dr. Xiaodong Liu
Project: Efficient Pretraining of Large Language Models based on ELECTRA.
Work submitted to ICML 2023.

Research Specialist @ Data Mining Lab, UCSD   July 2020 - June 2021
Advisor: Prof. Jingbo Shang
Project: Trustworthy Machine Learning

Research Assistant @ Intelligent Big Data Visualization Lab, Tongji University   July - Sept. 2019
Advisor: Prof. Nan Cao

Data Analyst Intern @ Research Center, Beijing Iminer Data Technology Co., ltd. July - Aug. 2018

Data Analyst Intern @ Data Evaluation Group, Baidu, Inc.   May - June 2018

Research Assistant @ Innovation Center for Data Science, SUSTech   Jan. - Mar. 2018
Advisor: Prof. Zhihong Xia

Talk	Deep-dive Panel at NeurIPS 2022	Dec. 2022
	Invited talk at Deep Learning Group, Microsoft Research	Sept. 2022
	The 37th International Conference on Machine Learning	July 2020
	International Symposium on Lunar and Planetary Science (ISLPS 2018), Macau	June 2018
	The 4th Symposium on Present and Future of Fundamental Astronomy, Hubei	June 2018

Service **Reviewer:** ICML 2021, NIPS 2021, ICLR 2022, ICML 2022, NIPS 2022, COLING 2022, ICLR 2023

Teaching	Teaching Assistant, Probability and Statistics for Data Science (DSE 210)	Winter 2023
	Teaching Assistant, Introduction to Machine Learning (CSE 151A)	Spring 2022
	Guest Speaker, Introduction to Machine Learning (CSE 151A)	Spring 2021

Award	Scholarship of National Astronomical Observatories	2015
	Academic Scholarship, 1st prize	2016, 2017
	Zhenggang Scholarship, top 40 graduate students in Nanjing University	2017
	Outstanding Student Leader, Nanjing University	2017