

Cheng Xu

Email: cheng.xu1@ucdconnect.ie Homepage: <https://xucheng.space/>

Last updated: September 25, 2024

EDUCATION	University College Dublin , Dublin, Ireland Ph.D., Computer Science Supervisor: Prof. M-Tahar Kechadi Thesis Title: Multimodal Fuzzy System for Fake News Detection Sep 2022 - Present
	Huaibei Normal University , Huaibei, P.R. China B.Eng., Data Science and Big Data Technology Mentor: Prof. Qi Fan & Prof. Longfeng Shen GPA: 90.34/100, Graduated with Provincial Honor (Highest for graduates) Sep 2018 - Jun 2022

PUBLICATIONS	* indicates corresponding author, † indicates co-first authorship, ‡ indicates a student I mentored.
	[10] Shuhao Guan, Cheng Xu , Moule Lin, Derek Greene. Effective Synthetic Data and Test-Time Adaptation for OCR Correction. (<i>EMNLP 2024</i>)
	[9] Yang Fan [‡] , Cheng Xu [*] , Shuhao Guan, Nan Yan, Yuke Mei. Advancing Arabic Sentiment Analysis: ArSen Benchmark and the Improved Fuzzy Deep Hybrid Network. (<i>CoNLL 2024</i>)
	[8] Cheng Xu [*] , Shuhao Guan, Derek Greene, M-Tahar Kechadi. Benchmark Data Contamination of Large Language Models: A Survey . (<i>Preprint, arXiv:2406.04244</i>)
	[7] Cheng Xu [*] , M-Tahar Kechadi. An Enhanced Fake News Detection System with Fuzzy Deep Learning . (<i>IEEE Access, 2024</i>)
	[6] Yang Fang [‡] , Cheng Xu . ArSen-20: A New Benchmark for Arabic Sentiment Detection . (<i>ICLR 2024, AfricaNLP Workshop</i>)
	[5] Nan Yan [†] , Cheng Xu . Decolonizing African NLP: A Comprehensive Survey on Power Dynamics and Data Colonialism in Tech Development . (<i>ICLR 2024, AfricaNLP Workshop</i>)
	[4] Cheng Xu , M-Tahar Kechadi. Fuzzy Deep Hybrid Network for Fake News Detection . (<i>SOICT 2023</i>)
	[3] Cheng Xu , Nan Yan [†] . AROT-COV23: A Dataset of 500K Original Arabic Tweets on COVID-19 . (<i>ICLR 2023, AfricaNLP Workshop</i>)
	[2] Cheng Xu , Jing Wang, Tianlong Zheng, Yue Cao, Fan Ye. Prediction of prognosis and survival of patients with gastric cancer by a weighted improved random forest model: an application of machine learning in medicine . (<i>Archives of Medical Science, 2022</i>)
[1] Cheng Xu , Qingling Chen, Fan Ye, Qi Fan, Qing Wang. Selection of surgical procedures and analysis of prognostic factors in patients with primary gastric tumour based on Cox regression: a SEER database analysis based on data mining . (<i>Gastroenterology Review, 2021</i>)	

HONORS	Graduate with Honor of Anhui Province (Top 3%) , Anhui Province Government 2022
	Graduate with Honor (Top 10%) , Huaibei Normal University 2022
	“Top 100 Outstanding College Students” honorary title, Anhui Province Government 2021
	“Top Ten College Students” honorary title, Huaibei Normal University 2021
	First Prize Scholarship for Innovation (Top 15) , Huaibei Normal University 2021
	First Prize Scholarship (Top 3%) Huaibei Normal University 2019 - 2021
Academic Excellence Award , Huaibei Normal University 2019 - 2021	

WORK EXPERIENCE	Ph.D. Researcher , Insight SFI Research Centre for Data Analytics Sep 2022 - Present
	▪ Received rigorous scientific training, emphasizing critical thinking and problem-solving skills.
	▪ Established a comprehensive false information recognition system capable of processing multimodal inputs, including news articles, textual context, and multimedia information, laying the groundwork for a robust and adaptable tool in the fight against the diffusion of fake news.
	▪ Spearheaded research efforts in leveraging deep learning architectures to tackle the multifaceted challenge of fake news detection, resulting in state-of-the-art performance on benchmark datasets such as LIAR.
	▪ Published findings in reputable academic venues and made code implementations publicly available, fostering collaboration and further advancements in the field of fake news detection and mitigation.

Research Assistant, Data Mining Lab of Huaibei Normal University Nov 2018 - Oct 2021

- Data mining and research on cancer data using machine learning techniques.
- Demonstrated leadership skills by leading and contributing to five research projects.
- Managed laboratory operations and resources in the third year of my undergraduate, enhancing my organizational and project management skills.

**PROJECT
EXPERIENCE**

COVID-19 Low-resource NLP Benchmark Development Jan 2023 - Present
Project Manager

- Managed a four-person team as Project Manager for the development of the COVID-19 Benchmark.
- Led the project resulting in the work like AROT-COV23 and Arsen-20, and published the papers at the AfricaNLP Workshop, a prestigious workshop in the field of low-resource NLP research.
- Currently overseeing benchmark testing, optimizing the dataset's utility for various NLP tasks.

Cancer Damage Prediction Using Data Mining and Machine Learning Sep 2020 - Jun 2022
Project Manager

- Led and supervised a cancer damage prediction project that was funded by the National University Student Innovation and Entrepreneurship Program in P.R. China, under [Grant 202010373032](#) and Grant S201910373149.
- Recommended the project achievements for participation in the 15th China University Students Innovation and Entrepreneurship Annual Conference.
- Spearheaded the entire development cycle of the project, including data preparation, collection, cleaning, imputation, modeling, and evaluation.
- Applied various machine learning models on the datasets, including both classic models (Random Forest, Decision tree) and our improved models.

TEACHING

Demonstrator, University College Dublin Sep 2023 - Present

- COMP40370 Data Mining (2023 Fall, 2024 Fall)
- COMP47780 Cloud Computing (2023 Fall, 2024 Fall)
- COMP10040 Introduction to Computer Architecture (2023 Fall, 2024 Fall)
- COMP10120 Computer Programming II (2024 Spring)
- COMP20230 Data Structure & Algorithms (Conv) (2024 Spring)
- COMP47980 Generative AI: Language Models (2024 Spring)
- COMP47590 Advanced Machine Learning (2024 Spring)
- COMP47650 Deep Learning (2024 Spring)
- COMP47580 Recommender Systems & Collective Intelligence (2024 Spring)
- COMP30760 Data Science in Python (2024 Fall)
- COMP30190 Program Construction II (2024 Fall)

**ACADEMIC
SERVER**

Conference Reviewer: NeurIPS, ICLR
Journal Reviewer: Archives of Medical Science