

CURRICULUM VITAE

Updated Feb 2024

PERSONAL INFORMATION

<i>Name:</i>	Linyi Yang
<i>Date of Birth:</i>	1995-07
<i>Mobile:</i>	+353(0)838557504 / +8617767072205
<i>Work Address:</i>	Engineering School, Westlake University, Hangzhou, Zhejiang, China
<i>Email:</i>	yanglinyi@westlake.edu.cn / yanglinyiucd@gmail.com
<i>Web Address:</i>	https://yanglinyi.github.io/
<i>Google Scholar</i>	https://scholar.google.com/citations?user=go3sFxcAAAAJ&hl=en&authuser=1
<i>Research Interest</i>	Large Language Models; Natural Language Processing; Causal Reasoning.
<i>Biography</i>	<p>Dr Linyi Yang is now working as a Research Assistant Professor at Westlake University. He has established academic reputation in the fusion of Natural Language Processing (NLP) and causal reasoning, alongside Large Language Models, as evidenced by 30+ publications, 13 first-author (co-first) publications in high-ranked venues (8 CCF A, 4 CCF B, 1 ICLR) with more than 1,200 citations, 1 national award, 1 best paper nomination award. He served as an Area Chair at EMNLP-22, Senior Program Committee Member at IJCAI-23. Additionally, he has successfully secured approximately £160K in funding and has mentored 4 PhD and over 10 graduate students as a subject adviser or co-supervisor.</p>

RESEARCH APPOINTMENTS

2023 - Now	Westlake University, China.	Research Assistant Professor
2021 - 2023	Westlake University, China. <ul style="list-style-type: none">- Topic: Casual-enhanced Deep Learning Methods for NLP.- Research Fund: CO-PI of “The New Generation of Pre-Trained Language Models”; 3M RMB- CO-PI of “Causal-enhanced Deep Learning Methods” funded by NSFC; 3M RMB- PI of “Causality-enhanced Methods for Building Robust NLP Models” funded by China Postdoctoral Science Foundation; 160K RMB	Research Associate
2020-2020	The University of Edinburgh, UK. <ul style="list-style-type: none">- Topic: Natural Language Processing in Finance- Host: Prof. Nan Yu	Visiting PhD Student

EDUCATION

Sep 2017 - Jul 2021	University College Dublin (UCD), Ireland. Ph.D. , Data Mining and Computer Science <ul style="list-style-type: none">- Thesis: Deep Neural Approach for Financial Analysis.- Supervisors: Prof Barry Smyth, Digital Chair of Computer Science, UCD; Fellow of the European Academy of Sciences.- <i>Outstanding Dissertation Award</i>	
Sep 2016 - Jul 2017	University College Dublin (UCD), Ireland. M.Sc. (Hons) Data Science. <ul style="list-style-type: none">- Thesis: Distant Supervised Relation Extraction (Published).- Supervisor: Dr Ruihai Dong.	(Master thesis: 90/100)
Sep 2012- Jul 2016	Harbin Engineering University, China. B.Sc. (Hons) Bachelor Software Engineering. <ul style="list-style-type: none">- Adviser: Prof Huiqiang Wang.	(Bachelor thesis: 90/100)

PROFESSIONAL EXPERIENCE

Area Chair	IJCAI-2023 (SPC) EMNLP-2022 (AC on Sentiment Analysis Track) CIKM-2022 (AC on Short Paper Track)
Editor	Special Issue of ACM Transactions on Information Systems on Evaluation of LLMs (with Jindong Wang, Sunayana Sitaram, Bhiksha Raj, and Qiang Yang)

HONORS AND AWARDS

2023	Outstanding Postdoctoral Representative: <i>Westlake University (1/180+)</i>
2022	Outstanding Self-Financed Overseas Students: <i>ONLY one Type-B winner in Ireland (1/600+)</i>
2022	Microsoft Research Asia Rising Star: <i>Host: Xiting Wang; Xing Xie</i>
2022	Outstanding Postdoctoral Researcher Awarded: <i>Westlake University</i>
2021	Outstanding Dissertation Award: <i>University College Dublin</i>
2020	Student Travel Grant: <i>The International World Wide Web Conference</i>
2019	Student Travel Grant: <i>The 28th International Joint Conference on Artificial Intelligence</i>
2018	Best Paper Nomination: <i>IEEE International Conference on Cloud Computing and Intelligence</i>
2017	Honor Class: <i>University College Dublin</i>

Teaching Experience

Traditional Teaching **Natural Language Processing (Guest lecturer: 2022 - 2023)**

Module: Natural Language Processing Level: PhD Students.

Textbook: Natural Language Processing: a machine learning perspective (Cambridge Press)

Natural Language Processing (Teacher Assistant)

Module: Natural Language Processing Level: PhD Students.

Oxford ML School (Teacher Assistant: 2021)

Module: Challenges in Natural Language Processing. Level: PhD and Industry Researchers.

Beijing-Dublin International College (BDIC Laboratory Lecturer: 2020)

Module: Databases and Info Systems. Level: Undergraduate Students.

Beijing-Dublin International College (BDIC Laboratory Lecturer: 2018 - 2019)

Module: Databases and Info Systems. Level: Undergraduate Students.

Module: Cloud Computing. Level: Undergraduate Students.

Student Supervision **PhD co-supervisor at Westlake University for Hongbo Zhang (2023)**

Thesis: Large Language Models for Logic Reasoning

PhD subject advisor at UCD for Yingjie Niu (2021)

Thesis: Deep Neural Approach for Analyzing Financial Documents from a Causal Perspective

PhD subject advisor at UCD for Xiao Li (2021)

Thesis: Optimization of Long-form Financial Document Processing

Co-supervisor at UCD for Several Bachelor Degree Projects (2018-2021)

Project 1: Deep Learning for Explainable Stock Prediction (Zheng Zhang, now PhD at Imperial College London).

Project 2: Bitcoin Trading System based on Reinforcement Learning (Michael Jordan, now master at UCD).

Project 3: Leveraging BERT for Improving FEARS Index for Financial Risk Prediction (Su Xiong, now master at NUS).

Project 4: Hierarchical Transformer for Stock Predictions (Lirui Wei, now master at UCL).

Project 5: Auto-grading Intelligent System (Zining Wang, now master at UCL).

Pedagogical approach **Educational outreach engagement – Organized by Insight Centre, UCD.**

It is becoming ever more important for scientists to be able to convey their research to a broad range of audiences. This module aims to develop skills in postgraduate students of science and technology related disciplines in a wide range of presentation styles and skills suitable for use in an outreach setting. Participants will be expected to engage in scholarly discussion with their peers on best practices from a communications and outreach perspective.

Invited Talks

Trustworthy Shanghai AI Lab, Shanghai, China 2023

Natural Language MLNLP, Online, China 2022

Processing MSRA, Beijing, China 2022

Causal-enhanced University of Manchester, UK 2023

Natural Language University of Warwick, UK 2022

Processing MSRA, Online, China 2022

Research Output

* indicates the equal contribution; h-index: 17; i10-index: 20, citations: 1,230 (up to Feb. 2024).

ICLR / NeurIPS/ ACL / WWW / AAAI are top-tier conferences.

Conference papers (C)

- [C1] **Linyi Yang**, Shuibai Zhang, Zhuohao Yu, Guangsheng Bao, Yidong Wang, Jindong Wang, Ruochen Xu, Wei Ye, Xing Xie, Weizhu Chen, Yue Zhang. Supervised Knowledge Makes Large Language Models Better In-context Learners. *ICLR 2024*, (2024). *(Collected by Hugging Face Daily Selection)*
- [C2] **Linyi Yang**, Yaoxiao Song, Xuan Ren, Chenyang Lyu, Yidong Wang, Lingqiao Liu, Jindong Wang, Jennifer Foster, Yue Zhang. Out-of-Distribution Generalization in Natural Language Processing: Past, Present, and Future. *EMNLP 2023, CCF B*, (2023).
- [C3] **Linyi Yang**, Yingpeng Ma, Yue Zhang. Measuring Consistency in Text-based Financial Forecasting Models, *Main Track, ACL 2023, CCF A*, (2023).
- [C4] **Linyi Yang**, Shuibai Zhang, Libo Qin, Yafu Li, Yidong Wang, Hanmeng Liu, Jindong Wang, Xing Xie, Yue Zhang. GLUE-X: Evaluating Natural Language Understanding Models from an Out-of-distribution Generalization Perspective, *ACL 2023 Findings, CCF A* (2023). *(Most cited paper at ACL-23 Findings. Reported by Many Media)*
- [C5] Yingjie Niu*, **Linyi Yang***, Ruihai Dong, Yue Zhang. Learning to Generalize for Cross-domain QA, *ACL 2023 Findings, CCF A* (2023).
- [C6] **Linyi Yang**, Lifan Yuan, Leyang Cui, Wenyang Gao, Yue Zhang. FactMix: Using a Few Labeled In-domain Examples to Generalize to Cross-domain Named Entity Recognition, *Main Track, COLING 2022, CCF B* (2022).
- [C7] **Linyi Yang**, Jiazheng Li, Ruihai Dong, Yue Zhang, Barry Smyth. NumHTML: Numeric-Oriented Hierarchical Transformer Model for Multi-task Financial Forecasting, *Main Track, AAAI 2022, CCF A* (2022).
- [C8] Jinghui Lu*, **Linyi Yang***, Brian Mac Namee, Yue Zhang. A Rationale-Centric Framework for Human-in-the-loop Machine Learning, *Main Track, ACL 2022, CCF A* (2022).
- [C9] **Linyi Yang**, Jiazheng Li, Pádraig Cunningham, Yue Zhang, Barry Smyth, Ruihai Dong. Exploring the Efficacy of Automatically Generated Counterfactuals for Sentiment Analysis, *Main Track, ACL 2021, CCF A* (2021).
- [C10] **Linyi Yang**, Eoin M Kenny, Tin Lok James Ng, Yi Yang, Barry Smyth, Ruihai Dong. Generating Plausible Counterfactual Explanations for Deep Transformers in Financial Text Classification, *Main Track, COLING 2020, CCF B* (2020).
- [C11] **Linyi Yang**, Tin Lok James Ng, Barry Smyth, Ruihai Dong. HTML: Hierarchical Transformer-based Multi-task Learning for Volatility Prediction, *Main Track, WWW 2020, CCF A* (2020).
- [C12] **Linyi Yang***, Jiazheng Li*, Barry Smyth, Ruihai Dong. MAEC: A Multimodal Aligned Earnings Conference Call Dataset for Financial Risk Prediction, *Main Track, CIKM 2020, CCF B* (2020).
- [C13] **Linyi Yang**, Yang Xu, Tin Lok James Ng, Ruihai Dong. Leveraging BERT to Improve the FEARS Index for Stock Forecasting, *IJCAI 2019, CCF A* (2019).
- [C14] **Linyi Yang**, Zheng Zhang, Su Xiong, Lirui Wei, Tin Lok James Ng, Lina Xu, Ruihai Dong. Explainable Text-Driven Neural Network for Stock Prediction, *Main Track, CCIS 2018, Best Paper Nomination* (2018).
- [C15] **Linyi Yang**, Tin Lok James Ng, Catherine Mooney, Ruihai Dong. Multi-level attention-based neural networks for distant supervised relation extraction, *Main Track, AICS 2017* (2017).
- [C16] **Linyi Yang**. Deep Neural Approach for Financial Analysis, *PhD Thesis, University College Dublin* (2021).
- [C17] Yidong Wang, Zhuohao Yu, Zhengran Zeng, **Linyi Yang**, Cunxiang Wang, Hao Chen, Chaoya Jiang, Rui Xie, Jindong Wang, Xing Xie, Wei Ye, Shikun Zhang, Yue Zhang. PandaLM: An Automatic Evaluation Benchmark for LLM Instruction Tuning Optimization. 2024, *ICLR 2024*. *(Hugging Face Monthly Downloads: 46,000+)*

- [C18] Guangsheng Bao, Yanbin Zhao, Zhiyang Teng, **Linyi Yang**, Yue Zhang. Fast-DetectGPT: Efficient Zero-Shot Detection of Machine-Generated Text via Conditional Probability Curvature. 2024, *ICLR 2024*. (340x faster than DetectGPT)
- [C19] Jindong Wang, Xixu Hu, Wenxin Hou, Hao Chen, Runkai Zheng, Yidong Wang, **Linyi Yang**, Haojun Huang, Wei Ye, Xiubo Geng, Binxin Jiao, Yue Zhang, Xing Xie. On the robustness of ChatGPT: An adversarial and out-of-distribution perspective. *ICLR 2023 Workshop RTML*, (2023). (*Most cited paper at ICLR-23 Workshop RTML*)
- [C20] Fang Guo, Yun Luo, **Linyi Yang**, Yue Zhang. SciMine: An Efficient Systematic Prioritization Model Based on Richer Semantic Information, *Main Track, SIGIR 2023, CCF A* (2023).
- [C21] Chenyang Lyu, **Linyi Yang**, Yue Zhang, Yvette Graham, Jennifer Foster. Exploiting Rich Textual User-Product Context for Improving Sentiment Analysis, *Main Track, ACL 2023, CCF A* (2023)
- [C22] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, **Linyi Yang**, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S Yu, Qiang Yang, Xing Xie. A survey on evaluation of large language models, *ACM Transactions on Intelligent Systems and Technology (TIST)*, (2023)
- [C23] Yile Wang, **Linyi Yang**, Zhiyang Teng, Ming Zhou, Yue Zhang. Pre-Training a Graph Recurrent Network for Language Representation. *NeurIPS 2nd Workshop on Efficient Natural Language and Speech Processing*, (2022).
- [C24] Yidong Wang, Hao Chen, Yue Fan, **Linyi Yang**, Bernt Schiele, Jindong Wang, Xing Xie, Yue Zhang and etc. USB: A Unified Semi-Supervised Learning Benchmark. *The Thirty-Sixth Annual Conference on Neural Information Processing Systems, NeurIPS 2022, CCF A* (2022).
- [C25] Yaoxian Song, Penglei Sun, Pengfei Fang, **Linyi Yang**, Yanghua Xiao and Yue Zhang. Human-in-the-loop Robotic Grasping using BERT Scene Representation, *Main Track, COLING 2022, CCF B* (2022).

Arxiv Papers (R)

- [R1] **Linyi Yang**, Zhen Wang, Yuxiang Wu, Jie Yang, Yue Zhang. Towards Fine-grained Causal Reasoning and QA. 2023
- [R2] Kaijie Zhu, Jindong Wang, Jiaheng Zhou, Zichen Wang, Hao Chen, Yidong Wang, **Linyi Yang**, Wei Ye, Neil Zhenqiang Gong, Yue Zhang, Xing Xie. PromptBench: Towards Evaluating the Robustness of Large Language Models on Adversarial Prompts. (Submission to ICML 2024)
- [R3] Yafu Li, Qintong Li, Leyang Cui, Wei Bi, Longyue Wang, **Linyi Yang**, Shuming Shi, Yue Zhang. Deepfake Text Detection in the Wild. (Submission to ACL 2024)
- [R4] Yun Luo, Hongjie Cai, **Linyi Yang**, Yanxia Qin, Rui Xia, Yue Zhang. Challenges for open-domain targeted sentiment analysis. 2022
- [R5] Yile Wang, **Linyi Yang**, Zhiyang Teng, Ming Zhou and Yue Zhang. Pre-Training a Graph Recurrent Network for Language Representation. *Transactions on Pattern Analysis and Machine Intelligence*. (TPAMI; Under Revision)

Thesis and Technical Reports (T)

- [T1] **Linyi Yang**. Deep Neural Approach for Financial Analysis. PhD Thesis, UCD. (2021)

Funding History

- [F1] **CO-PI** of “Causal-enhanced Deep Learning Methods” (2023-2026), **3M RMB**, funded by NSFC
- [F2] **CO-PI** of “The new generation of pre-trained language models” (2022-2025), **3M RMB**, funded by Zhejiang Province Key Project
- [F3] **PI** of “Causality-enhanced Methods for Building Robust NLP Models” (2022), **160K RMB**, funded by China Postdoctoral Science Foundation
- [F4] **Winner** of “Deep neural approach for financial analysis” (2017-2021), **€ 124,400**, full PhD scholarship funded by Science Foundation Ireland (SFI)
- [F5] **Winner** of “Outstanding Self-Funded Foreign Students Candidate” (2021), **10K Dollars**, funded by NSFC