

## Linyi Yang, Ph.D.

---

CONTACT INFORMATION	Westlake University, Hangzhou, China. Work: +86 17767072205 Email: yanglinyi@westlake.edu.cn	Google Scholar; GitHub Website: yanglinyi.github.io
RESEARCH INTERESTS	Causality; Trustworthy AI; Data Science; Applications of Natural Language Processing.	
RESEARCH APPOINTMENTS	<b>Westlake University</b> , Hangzhou, China Postdoctoral Research Fellow - Topic: <i>Casual-enhanced Deep Learning Methods for NLP</i> - Co-PI: <i>The New Generation of Pre-trained Language Models</i> , 3M RMB - Co-PI: <i>Causal-enhanced Deep Learning Methods</i> , NSFC Key Project, 3M RMB - PI: <i>Causal-enhanced Methods for Building Robust NLP Models</i> , China Postdoctoral Science Foundation, 160K RMB	2021-Present
	<b>The University of Edinburgh</b> , UK Visiting PhD Student - Topic: <i>Natural Language Processing in Finance</i> - Host: <i>Prof. Nan Yu</i>	2020-2020
EDUCATION	<b>University College Dublin (UCD)</b> , Ireland <b>Ph.D.</b> , Data Mining and Computer Science - Topic: <i>Deep Neural Approach for Financial Analysis</i> - Advisor: <i>Prof. Barry Smyth</i> (Fellow of the European Academy of Sciences, Top 20 Highest Cited Scientist in Recommender System) - <i>Outstanding Dissertation Award</i>	2017-2021
	<b>University College Dublin (UCD)</b> , Ireland <b>M.Sc.</b> , (Hons) Data Science - Topic: <i>Distant Supervised Relation Extraction</i> - Advisor: <i>Dr Ruihai Dong</i>	2016-2017
	<b>Harbin Engineering University</b> , China <b>B.Sc.</b> , (Hons) Bachelor Software Engineering - Advisor: <i>Prof Huiqiang Wang</i>	2012-2016
PROFESSIONAL SERVICE	<b>Academic Leadership</b> - <b>Senior PC</b> : International Joint Conference on Artificial Intelligence (IJCAI) - <b>Area Chair</b> : Empirical Methods in Natural Language Processing (EMNLP)	2023 2022
HONORS AND AWARDS	- <b>Outstanding Postdoctoral Representative</b> : ONLY awardee Last year: Dr Rui Bai - <b>Microsoft Early-Career Rising Star</b> : MSRA Host: Xiting Wang, Xing Xie - <b>Outstanding Self-Financed Students Abroad</b> : ONLY one Type-B winner in Ireland - Outstanding Postdoctoral Award: Westlake University - Outstanding Dissertation Award: University College Dublin - Student Travel Grant: The International World Wide Web Conference - Student Travel Grant: The International Joint Conference on Artificial Intelligence - <b>Best Paper Nomination</b> : IEEE International Conference on Cloud Computing - Honor Class: University College Dublin - Honor Class: Harbin Engineering University	2023 2022 2022 2022 2021 2020 2019 2018 2017 2016

## PUBLICATIONS

\* indicates the equal contribution. [ACL/NeurIPS/WWW/IJCAI/AAAI](#) are top-tier conferences

- [1] **Linyi Yang**, Yingpeng Ma, Yue Zhang. Measuring Consistency in Text-based Financial Forecasting Models, *Main Track, ACL 2023, Core A\**, 2023.
- [2] **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, Core A\**, 2023 *Most cited paper at ACL-23. Reported by many Media.*
- [3] **Linyi Yang\***, Yingjie Niu\*, Ruihai Dong, Yue Zhang. Learning to Generalize for Cross-domain QA, *ACL 2023 Findings, Core A\**, 2023.
- [4] **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, *COLING 2022, Core A*, 2022.
- [5] **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, Core A\**, 2022.
- [6] **Linyi Yang\***, Jinghui Lu\*, Brian Mac Namee, Yue Zhang. A Rationale-Centric Framework for Human-in-the-loop Machine Learning, *Main Track, ACL 2022, Core A\**, 2022.
- [7] **Linyi Yang**, Jiazheng Li, Pdraig Cunningham, Yue Zhang, Barry Smyth, Ruihai Dong. Exploring the Efficacy of Automatically Generated Counterfactuals for Sentiment Analysis, *Main Track, ACL 2021, Core A\**, 2021.
- [8] **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, Core A*, 2020.
- [9] **Linyi Yang**, Tin Lok James Ng, Barry Smyth, Ruihai Dong. HTML: Hierarchical Transformer-based Multi-task Learning for Volatility Prediction, *Main Track, WWW 2020, Core A\**, 2020.
- [10] **Linyi Yang\***, Jiazheng Li\*, Barry Smyth, Ruihai Dong. MAEC: A Multimodal Aligned Earnings Conference Call Dataset for Financial Risk Prediction, *Main Track, CIKM 2020, Core A*, 2020.
- [11] **Linyi Yang**, Yang Xu, Tin Lok James Ng, Ruihai Dong. Leveraging BERT to Improve the FEARS Index for Stock Forecasting, *IJCAI 2019, Core A\**, 2019.
- [12] **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.
- [13] **Linyi Yang**, Tin Lok James Ng, Catherine Mooney, Ruihai Dong. Multi-level attention-based neural networks for distantly supervised relation extraction, *Main Track, AICS 2017*.
- [14] **Linyi Yang**. Deep Neural Approach for Financial Analysis, *PhD Thesis, University College Dublin*, 2021.
- [15] 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*.

- [16] Fang Guo, Yun Luo, **Linyi Yang**, Yue Zhang. SciMine: An Efficient Systematic Prioritization Model Based on Richer Semantic Information, *Main Track, SIGIR 2023*.
- [17] Chenyang Lyu, **Linyi Yang**, Yue Zhang, Yvette Graham, Jennifer Foster. Exploiting Rich Textual User-Product Context for Improving Sentiment Analysis, *ACL 2023 Findings*.
- [18] 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. *A Trending Paper on Twitter*.
- [19] 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.
- [20] Yidong Wang, Hao Chen, Yue Fan, Wang Sun, Ran Tao, Wenxin Hou, Renjie Wang, **Linyi Yang**, Zhi Zhou, Lan-Zhe Guo, Heli Qi, Zhen Wu, Yu-Feng Li, Satoshi Nakamura, Wei Ye, Marios Savvides, Bhiksha Raj, Takahiro Shinozaki, Bernt Schiele, Jindong Wang, Xing Xie, Yue Zhang. USB: A Unified Semi-supervised Learning Benchmark for Classification, *The Thirty-Sixth Annual Conference on Neural Information Processing Systems, NeurIPS 2022*.
- [21] 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*.
- ARXIV PAPERS
- [22] **Linyi Yang**, Yaoxiao Song, Xuan Ren, Chenyang Lyu, Yidong Wang, Lingqiao Liu, Jindong Wang, Jennifer Foster, Yue Zhang. Out-of-Distribution Generalization in Text Classification: Past, Present, and Future, Arxiv, 2023.
- [23] **Linyi Yang**, Zhen Wang, Yuxiang Wu, Jie Yang, Yue Zhang. Towards Fine-grained Causal Reasoning and QA, Arxiv, 2023.
- [24] 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, Arxiv, 2023, *Hugging Face Monthly Downloads: 48,000+*.
- [25] 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, Arxiv, 2023, *Reported by many Media*.
- [26] Yafu Li, Qintong Li, Leyang Cui, Wei Bi, Longyue Wang, **Linyi Yang**, Shuming Shi, Yue Zhang. Deepfake Text Detection in the Wild, Arxiv, 2023.
- [27] Yun Luo, Hongjie Cai, **Linyi Yang**, Yanxia Qin, Rui Xia, Yue Zhang. Challenges for open-domain targeted sentiment analysis, Arxiv, 2023.
- [28] Yile Wang, **Linyi Yang**, Zhiyang Teng, Ming Zhou and Yue Zhang. Pre-Training a Graph Recurrent Network for Language Representation, Arxiv, 2023. (Submission to Transactions on Pattern Analysis and Machine Intelligence)
- [29] **Linyi Yang**, TLJ Ng, B Smyth, R Dong. Fact Check: Analyzing Financial Events from Multilingual News Sources. Arxiv, 2021.
- [30] **Linyi Yang**. Distant Supervised Relation Extraction. M.Sc. Thesis. Arxiv, 2017.

SELECTED OPEN-SOURCE TOOLS	<p>★ represents the number of github stars. # represents my intern's work.</p> <ul style="list-style-type: none"> <li>- # <b>USB</b> ★ 880 <a href="https://github.com/microsoft/Semi-supervised-learning">https://github.com/microsoft/Semi-supervised-learning</a> USB: A Unified Semi-supervised learning Benchmark for CV, NLP, and Audio Classification, which provides the implementation of 14 SSL algorithms based on Consistency Regularization, and 15 tasks for evaluation from the CV, NLP, and Audio domain.</li> <li>- <b>FinNLP</b> ★ 360 <a href="https://github.com/YangLinyi/FinNLP-Progress">https://github.com/YangLinyi/FinNLP-Progress</a> A repository to track the progress in Natural Language Processing (NLP) related to the domain of Finance, including the datasets, papers, and current state-of-the-art algorithms.</li> <li>- <b>GLUE-X</b> ★ 120 <a href="https://github.com/YangLinyi/GLUE-X">https://github.com/YangLinyi/GLUE-X</a> We leverage 14 datasets as OOD test data and conduct evaluations on 8 NLU tasks over 21 popularly used models.</li> <li>- # <b>PandaLM</b> ★ 660 <a href="https://github.com/WeOpenML/PandaLM">https://github.com/WeOpenML/PandaLM</a> PandaLM aims to provide reproducible and automated comparisons between different LLMs. It can compare the responses of different LLMs and provide a reason for the decision, along with a reference answer. .</li> </ul>
INVITED TALKS	<p><b>Trustworthy Natural Language Processing</b></p> <ul style="list-style-type: none"> <li>- Shanghai AI Lab, Shanghai, China (03/2023)</li> <li>- MLNLP, Online, China (11/2022)</li> <li>- MSRA, Beijing, China (10/2022)</li> </ul> <p><b>Causal-enhanced Natural Language Processing</b></p> <ul style="list-style-type: none"> <li>- University of Manchester, UK (10/2022)</li> <li>- University of Warwick, UK (08/2022)</li> <li>- MSRA, Online, China (07/2022)</li> </ul>
MENTORED STUDENTS	<p><b>Westlake University</b> China</p> <p>Yidong Wang (Master Student) 2021-2023</p> <ul style="list-style-type: none"> <li>- PandaLM Submitted to NeurIPS 2023</li> <li>- USB NeurIPS 2022</li> <li>- Freematch ICLR 2023</li> <li>- On the Robustness of ChatGPT ICLR 2023 Workshop</li> </ul> <p>Yichun Qian (Undergraduate Student from Zhejiang University) 2022-2023</p> <ul style="list-style-type: none"> <li>- Now Master Student at Stanford University</li> </ul> <p>Shuibai Zhang (Undergraduate Student from UESTC) 2022-2023</p> <ul style="list-style-type: none"> <li>- GLUE-X ACL 2023</li> </ul> <p>Yingpeng Ma (UG from Northwestern Polytechnical University) 2022-2023</p> <ul style="list-style-type: none"> <li>- FinTrust ACL 2023</li> </ul> <p>Fang Guo (Ph.D. from Westlake University) 2022-2023</p> <ul style="list-style-type: none"> <li>- SciMine SIGIR 2023</li> </ul> <p><b>UCD</b> Ireland</p> <p>Jiazheng Li (Undergraduate Student) 2020-2022</p> <ul style="list-style-type: none"> <li>- NumHTML AAI 2022</li> <li>- Counterfactual Generation ACL 2021</li> <li>- MAEC CIKM 2020</li> </ul> <p>Yingjie Niu (Master Student) 2022</p> <ul style="list-style-type: none"> <li>- Cross-domain QA ACL 2023</li> </ul>
TEACHING EXPERIENCE	<p><b>Westlake University, China</b></p> <ul style="list-style-type: none"> <li>- Natural Language Processing in English. (TA for PhD Students) 2022-2023</li> </ul> <p><b>Oxford ML School, UK</b></p> <ul style="list-style-type: none"> <li>- The Challenge of Natural Language Processing in English. (TA for Graduate Students) 2021</li> </ul> <p><b>Beijing-Dublin International College, Ireland</b></p> <ul style="list-style-type: none"> <li>- Introduction to Programming in English. (Laboratory Lecturer) 2018-2020</li> <li>- Cloud Computing in English. (Laboratory Lecturer) 2018-2020</li> <li>- Database System in English. (Laboratory Lecturer) 2018-2020</li> </ul>

REFERENCES

**Yue Zhang** (Postdoc Advisor)

*Professor*

Engineering School, Westlake University, China

zhangyue@westlake.edu.cn

**Barry Smyth** (Doctoral Advisor)

*Chair Professor*

Digital Chair of Computer Science, Insight Centre for Data Analytics, UCD

barry.smyth@ucd.ie

**Jie Yang**

*Assistant Professor*

Harvard Medical School, Harvard University, USA

jjeynlp@gmail.com

**Xing Xie**

*Senior Principal Research Manager*

Microsoft Research Asia, China

xingx@microsoft.com

**Jindong Wang**

*Senior Researcher*

Microsoft Research Asia, China

jindwang@microsoft.com

FUNDING HISTORY

**CO-PI of “Causal-enhanced Deep Learning Methods”,**

- 3 Million RMB, National Key Project (3/28)

2024-2029

**CO-PI of “The new generation of pre-trained language models”,**

- 3 Million RMB, Zhejiang Province Key Project (2/16)

2023-2026

**PI of “Causality-enhanced Methods for Building Robust NLP Models”,**

- 160K RMB, China Postdoctoral Science Foundation

2022

**Scholarship of “Deep neural approach for financial analysis”,**

- 125K Euros, Full Scholarship Funded by European Union

2017-2021

**Winner of “Outstanding Self-Funded Foreign Students Candidate” ,**

- 10K Dollars, China Scholarship Council

2022