Yi Yang

CONTACT
INFORMATION

Department of Information Systems, Business Statistics, and Operations Management,
School of Business and Management,
Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong   imyiyang@ust.hk

EMPLOYMENT

Hong Kong University of Science and Technology

Associate Professor,   July 2023-present
School of Business and Management

Assistant Professor,   July 2017-June 2023
School of Business and Management

Affiliated Faculty,   July 2019-present
Risk Management and Business Intelligence Program

EDUCATION

Northwestern University, Evanston, Illinois
Ph.D. in Computer Science   Aug 2015

University of Science and Technology of China, Anhui, China
B.S. in Computer Science   July 2009

RESEARCH INTERESTS

Method Area: Machine Learning, Natural Language Processing, Deep Learning
Applied Area: Financial Technology, Marketing Analytics, Network Analytics

JOURNAL
PUBLICATIONS

1. John P Lalor, Ahmed Abbasi, Kezia Oketch, **Yi Yang**, and Nicole Forsgren. Should fairness be a metric or a model? a model-based framework for assessing bias in machine learning pipelines. *ACM Transactions on Information Systems*, 2024

2. Dongcheng Zhang, Kunpeng Zhang, **Yi Yang**, and David Schweidel. Tm-okc: An unsupervised topic model for text in online knowledge communities. *Accepted at MIS Quarterly*

3. **Yi Yang** and Ramanath Subramanyam. Extracting actionable insights from text data: a stable topic model approach. *MIS Quarterly*, 47(3), 2023

4. Allen H Huang, Hui Wang, and **Yi Yang**. Finbert: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 40(2):806–841, 2023

5. **Yi Yang**, Yu Qin, Yangyang Fan, and Zhongju Zhang. Unlocking the power of voice for financial risk prediction: A theory-driven deep learning design approach. *MIS Quarterly*, 47(1), 2023

6. **Yi Yang**, Kunpeng Zhang, and Yangyang Fan. sdtm: A supervised bayesian deep topic model for text analytics. *Information Systems Research*, 34(1):137–156, 2023

7. Zhuoyi Peng, **Yi Yang**, Liu Yang, and Kai Chen. Federated meta embedding concept stock recommendation. *IEEE Transactions on Big Data*, 2022

8. Zhe Tang, **Yi Yang**, Wen Li, Defu Lian, and Lixin Duan. Deep cross-attention network for crowdfunding success prediction. *IEEE Transactions on Multimedia*, 2022

9. **Yi Yang**, Kunpeng Zhang, and PK Kannan. Identifying market structure: A deep network representation learning of social engagement. *Journal of Marketing*, 86(4):37–56, 2022

10. **Yi Yang**, Kunpeng Zhang, and Yangyang Fan. Analyzing firm reports for volatility prediction: A knowledge-driven text-embedding approach. ***INFORMS Journal on Computing***, 34(1):522–540, 2022

11. Fan Zhou, Kunpeng Zhang, Bangying Wu, **Yi Yang**, and Harry Jiannan Wang. Unifying online and offline preference for social link prediction. ***INFORMS Journal on Computing***, 33(4):1400–1418, 2021

12. **Yi Yang**, Shimei Pan, Jie Lu, Mercan Topkara, and Yangqiu Song. The stability and usability of statistical topic models. ***ACM Transactions on Interactive Intelligent Systems (TiiS)***, 6(2):1–23, 2016

13. Tiantian Zhu, Hongyu Gao, **Yi Yang**, Kai Bu, Yan Chen, Doug Downey, Kathy Lee, and Alok N Choudhary. Beating the artificial chaos: Fighting osn spam using its own templates. ***IEEE/ACM Transactions on Networking***, 24(6):3856–3869, 2016

14. Fei Huang, Arun Ahuja, Doug Downey, **Yi Yang**, Yuhong Guo, and Alexander Yates. Learning representations for weakly supervised natural language processing tasks. ***Computational Linguistics***, 40(1):85–120, 2014

<small>CONFERENCE PUBLICATIONS</small>

1. Yue Guo, Chenxi Hu, and **Yi Yang**. Predict the future from the past? on the temporal data distribution shift in financial sentiment classifications. In ***Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing***, pages 1029–1038, 2023

2. Yue Guo, Zian Xu, and **Yi Yang**. Is chatgpt a financial expert? evaluating language models on financial natural language processing. In ***Findings of the Association for Computational Linguistics: EMNLP 2023***, pages 815–821, 2023

3. Yixuan Tang, **Yi Yang**, Allen Huang, Andy Tam, and Justin Tang. Finentity: Entity-level sentiment classification for financial texts. In ***Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing***, pages 15465–15471, 2023

4. Fan Zhou, Yuzhou Mao, Liu Yu, **Yi Yang**, and Ting Zhong. Causal-debias: Unifying debiasing in pretrained language models and fine-tuning via causal invariant learning. In ***Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics***, pages 4227–4241, 2023

5. Yue Guo, **Yi Yang**, and Ahmed Abbasi. Auto-debias: Debiasing masked language models with automated biased prompts. In ***Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics***, pages 1012–1023, 2022

6. Chengyu Chuang and **Yi Yang**. Buy tesla, sell ford: Assessing implicit stock market preference in pre-trained language models. In ***Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics***, pages 100–105, 2022

7. John P Lalor, **Yi Yang**, Kendall Smith, Nicole Forsgren, and Ahmed Abbasi. Benchmarking intersectional biases in nlp. In ***Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies***, pages 3598–3609, 2022

8. **Yi Yang**, Chewei Liu, and Sunil Mithas. Dual information technology strategy and firm performance: Insights from a deep learning approach. In ***Academy of Management Proceedings***, volume 2021, page 12694. Academy of Management Briarcliff Manor, NY 10510, 2021

9. Zhenyu He, Ce Li, Fan Zhou, and **Yi Yang**. Rumor detection on social media with event augmentations. In ***Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval***, pages 2020–2024, 2021

10. Ahmed Abbasi, David Dobolyi, John P Lalor, Richard G Netemeyer, Kendall Smith, and **Yi Yang**. Constructing a psychometric testbed for fair natural language processing. In ***Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing***, pages 3748–3758, 2021

11. Hanyu Duan, **Yi Yang**, and Kar Yan Tam. Learning numeracy: A simple yet effective number embedding approach using knowledge graph. In ***Findings of the Association for Computational Linguistics: EMNLP 2021***, pages 2597–2602, 2021

12. Linyi Yang, Eoin Kenny, Tin Lok James Ng, **Yi Yang**, Barry Smyth, and Ruihai Dong. Generating plausible counterfactual explanations for deep transformers in financial text classification. In ***Proceedings of the 28th International Conference on Computational Linguistics***, pages 6150–6160, 2020

13. Xinyi Wang and **Yi Yang**. Neural topic model with attention for supervised learning. In ***International Conference on Artificial Intelligence and Statistics***, pages 1147–1156. PMLR, 2020

14. Ting Zhong, Tianliang Wang, Fan Zhou, Goce Trajcevski, Kunpeng Zhang, and **Yi Yang**. Interpreting twitter user geolocation. In ***Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics***, pages 853–859, 2020

15. Fan Zhou, Shengming Zhang, and **Yi Yang**. Interpretable operational risk classification with semi-supervised variational autoencoder. In ***Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics***, pages 846–852, 2020

16. Yu Qin and **Yi Yang**. What you say and how you say it matters: Predicting stock volatility using verbal and vocal cues. In ***Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics***, pages 390–401, 2019

17. Fan Zhou, Bangying Wu, **Yi Yang**, Goce Trajcevski, Kunpeng Zhang, and Ting Zhong. Vec2link: Unifying heterogeneous data for social link prediction. In ***Proceedings of the 27th ACM International Conference on Information and Knowledge Management***, pages 1843–1846, 2018

18. Doug Downey, Chandra Bhagavatula, and **Yi Yang**. Efficient methods for inferring large sparse topic hierarchies. In ***Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)***, pages 774–784, 2015

19. **Yi Yang**, Shimei Pan, Yangqiu Song, Jie Lu, Mercan Topkara, and JW Player. Improving topic model stability for effective document exploration. In ***IJCAI***, pages 4223–4227, 2016

20. **Yi Yang**, Doug Downey, and Jordan Boyd-Graber. Efficient methods for incorporating knowledge into topic models. In ***Proceeding of Empirical Methods in Natural Language Processing***, 2015

21. **Yi Yang**, Shimei Pan, Yangqiu Song, Jie Lu, and Mercan Topkara. User-directed non-disruptive topic model update for effective exploration of dynamic content. In ***Proceedings of the 20th International Conference on Intelligent User Interfaces***, pages 158–168, 2015

22. Kunpeng Zhang, **Yi Yang**, Aaron Sun, and Hengchang Liu. A systematic framework for sentiment identification by modeling user social effects. In ***2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)***, volume 2, pages 172–179. IEEE, 2014

23. Hongyu Gao, **Yi Yang**, Kai Bu, Yan Chen, Doug Downey, Kathy Lee, and Alok Choudhary. Spam ain't as diverse as it seems: throttling osn spam with templates underneath. In ***Proceedings of the 30th Annual Computer Security Applications Conference***, pages 76–85, 2014

24. **Yi Yang**, Alexander Yates, and Doug Downey. Overcoming the memory bottleneck in distributed training of latent variable models of text. In ***Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies***, pages 579–584, 2013

GRANTS

Principal Investigator (PI)

1. HKUST-WeBank Joint Lab Project, HKD $497,000, 2022-2023
   Title: Pre-training for Tabular Business Data

2. HKUST-Kaisa Joint Research Institute Seed Fund, HKD $670,000, 2021-2022
   Title: Deep Learning in Financial NLP

3. Hong Kong Research Grants Council, Early Career Scheme, HKD $458,000, 2019-2022
   Title: Beyond Text: The Information Role of Voice in Financial Market

4. University Initial Grant, HKD $100,000, 2017-2020

Co-Investigator (Co-I)

1. Hong Kong Research Grants Council, Theme-based Research Scheme, HKD $20,659,000, 2018-2022
   Title: Contributing to the Development of Hong Kong into a Global Fintech Hub

2. National Natural Science Foundation of China, RMB $2,570,000, 2020-2023
   Title: Technologies of Fusion Perception and Collaborative Processing for Behavior Big Data in Cyberspace

3. Hong Kong Research Grants Council, General Research Fund, HKD $605,216, 2022-2023
   Title: Product Recommendations in Email Campaigns for Passive Customers

WORK EXPERIENCE

University of Illinois Urbana-Champaign, Urbana, Illinois
Postdoc Research Associate, College of Business   Aug 2015-June 2017

Amazon, Seattle, Washington
Software Engineer Intern, Amazon Web Services   Sep 2013-Dec 2013

IBM T.J. Watson Research, Yorktown Heights, New York
Research Intern, Intelligent Multimedia Interaction Division   May 2013-Sep 2013

TEACHING

Hong Kong University of Science and Technology ,
Instructor,   July 2017-present

    ISOM3360: Data Mining for Business Analytics, (Undergradaute)
    ISOM3370: Big Data Technologies, (Undergradaute)
    ISOM5270: Big Data Analytics (MBA full-time and part-time)
    ISOM5270: Big Data Analytics (Master)

★ Finalist for *The Franklin Prize for Teaching Excellence* for Year 2021
★ Finalist for *The Franklin Prize for Teaching Excellence* for Year 2018

University of Illinois Urbana-Champaign, Urbana, Illinois
Instructor,   Aug 2015-June 2017

    BADM352: Database Design and Management

★ Awarded *Teachers Ranked as Excellent by Their Students* in Spring 2017 and Fall 2016.

STUDENTS
SUPERVISION

Postgraduate students:
- Yingpeng Zhu, PhD candidate, ISOM, Supervisor, 2017-2023 (expected)
- Jiaxin Liu, PhD candidate, FinTech, Co-supervisor (with Prof. Kar Yan Tam at ISOM), 2020-present
- Hanyu Duan, MPhil, ISOM, Co-supervisor (with Prof. Kar Yan Tam at ISOM), 2021-present
- Yue Guo, PhD Candidate DSA, Co-supervisor (with Prof. Yangqiu Song at CSE) 2020-present
- Zihan Meng, MPhil, CSE, Co-supervisor (with Prof. Qiong Luo at CSE) 2019-2021

Undergraduate students:
- 30 undergraduate students that participated in the Undergraduate Research Opportunities Program (UROP) at HKUST, July 2017- present

SERVICE

University Service

PhD/MPhil Committee (IS), 2023-2024
HKUST JUPAS Interview Taster class, 2023
FT/PT MScISM Committee (Information Systems), 2022-2024
Search and Appointments Committee (Information Systems), 2021-2022
UG committee (Information Systems), 2020-2021
Technology Review Committee (TRC), 2022-present

Editorial Board

Associate Editor, Information Technology and Management (2020-present)

External Reviewer

Information Systems Research, Management Information Systems Quarterly, Management Science, INFORMS Journal on Computing, Production and Operations Management, Decision Support Systems, Journal of the Association for Information Systems