Yaqing Wang

wang5075@purdue.edu | yaqingwang | (+1)858-766-0845
yaqingwang.github.io | Google Scholar

EDUCATION

Purdue University — West Lafayette, IN, USA
Ph.D. in Electrical and Computer Engineering — 2020 - Present
Advisor: Prof. Jing Gao

University at Buffalo — Buffalo, NY, USA
Ph.D. in Computer Science and Engineering — 2016 - 2020
Advisor: Prof. Jing Gao

University of California San Diego — La Jolla, CA, USA
M.Sc. in Statistics — 2014 - 2016

Shandong University — Jinan, Shandong, China
BSc in Mathematics — 2009 - 2013

RESEARCH INTERESTS

I am broadly interested in data science and artificial intelligence with a focus on data mining and machine learning. In particular, I am interested in Data-efficient Learning, Information Trustworthiness Evaluation and Knowledge Discovery.

RESEARCH EXPERIENCE

Data Mining and Machine Learning Lab — Sep. 2016 - Present
Research Assistant — Advisor: Jing Gao

Project i: *Data Efficient Learning*

- A major bottleneck of deep learning models is the reliance on massive sets of hand-labeled training data. We aim to improve data efficiency for deep learning models when labeled data is scarce or noisy. The explorations are in the meta-learning [KDD'20a, KDD'21], weak supervision [AAAI'20], few-shot/zero-shot learning [KDD'21, 24], and data augmentation [SDM'20].

Project ii: *Information Trustworthiness Evaluation*

- The explosion of information from a variety of sources has made it increasingly important to check the credibility and reliability of the data . We devote the efforts to mitigate misinformation [KDD'18a, AAAI'20, KDD'21] and validate the extracted knowledge [CIKM'20, KDD'20a].

Project iiii: *Knowledge Discovery*

- Lots of human knowledge is encoded in text. To make knowledge resources more findable, accessible, interoperable, and reusable (FAIR), we focus on extracting structured knowledge from massive collection of text [KDD'21, KDD'20b, Bioinfomatics].

LIA Group @ Microsoft Research Redmond — May-Aug 2021
Research Intern

- Mentor: Dr. Subhabrata Mukherjee, Dr. Xiaodong Liu, Dr. Ahmed Hassan Awadallah and Dr. Jianfeng Gao

LIT Group @ Microsoft Research Redmond — May-Aug 2020
Research Intern

– Mentors: Dr. Subhabrata Mukherjee and Dr. Ahmed Hassan Awadallah.
– One paper [24] at KDD 2021

Product Graph Group @ Amazon, Seattle May-Aug 2019
Applied Scientist Intern

– Mentors: Dr. Xin Luna Dong, Dr. Xian Li and Dr. Yifan Ethan Xu.

– Two papers [14, 15] at KDD 2020

PROFESSIONAL SERVICE

Journal Reviewer IEEE Access, VLDB, WIREs Data Mining and Knowledge Discovery, ACM Computing Surveys (CSUR), TKDD, IEEE Transactions on Multimedia, World Wide Web Journal

PC Member ICML (2020, 2021), KDD (2020, 2021), Big Data (2020), AAAI (2021), ICLR (2021), NAACL (2021), ACL (2021), EMNLP (2021), NeurIPS (2021)

Conference Volunteer KDD 2020, KDD 2018, ICDM 2017

PUBLICATIONS

Preprints & Submissions

[24] *Learning from Language: Low-shot Named Entity Recognition via Decomposed Framework*
Yaqing Wang, Haoda Chu, Chao Zhang, Jing Gao.
In submission.

[23] *AIM: Adversarial Inference by Matching Priors and Conditionals.*
Hanbo Li*, **Yaqing Wang***, Changyou Chen, Jing Gao. (* Equal Contribution)

[22] *FedSemi: An Adaptive Federated Semi-Supervised Learning Framework*
Zewei Long, Liwei Che, **Yaqing Wang**, Muchao Ye, Junyu Luo, Jinze Wu, Houping Xiao, Fenglong Ma.

[21] *MedLane: A Benchmark Dataset for Understandable Medical Language Translation*
Junyu Luo, Zifei Zheng, Hanzhong Ye, Muchao Ye, **Yaqing Wang**, Quanzeng You, Cao Xiao, Fenglong Ma.

Peer-Reviewed Conference and Journal Papers (* equal contribution)

[20] *Meta Self-training for Few-shot Neural Seuqence Labeling*
Yaqing Wang, Subhabrata Mukherjee, Haoda Chu, Yuancheng Tu, Ming Wu, Jing Gao, Ahmed Hassan Awadallah.
Proceedings of 2021 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2021**), August 2021
Acceptance rate: 238/1541=15.4%

[19] *Multimodal Emergent Fake News Detection via Meta Neural Process Networks*
Yaqing Wang, Fenglong Ma, Haoyu Wang, Kishlay Jha and Jing Gao.
Proceedings of 2021 ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2021**), August 2021
Acceptance rate: 138/705=19.6%

[18] *MedPath: Augmenting Health Risk Prediction via Medical Knowledge Paths.*
Muchao Ye, Suhan Cui, **Yaqing Wang**, Junyu Luo, Cao Xiao and Fenglong Ma.
Proceedings of the 30th The Web Conference (**WWW 2021**), Ljubljana, Slovenia, April 19-23, 2021.
Acceptance rate: 357/1736=20.6%

[17] *Fair Classification Under Strict Unawareness.*
Haoyu Wang, Hengtong Zhang, **Yaqing Wang** and Jing Gao.
Proceedings of the SIAM International Conference on Data Mining (**SDM 2021**), March 25 - 27, 2021, Alexandria, Virginia, US.
Acceptance rate: 85/400=21.25%

[16] *Towards Learning Outcome Prediction via Modeling Question Explanations and Student Responses.*
Tianqi Wang, Fenglong Ma, **Yaqing Wang**, Tang Tang, Longfei Zhang, and Jing Gao.
Proceedings of the SIAM International Conference on Data Mining (**SDM 2021**), March 25 - 27, 2021, Alexandria, Virginia, US.
Acceptance rate: 85/400=21.25%

[15] *Automatic Validation of Textual Attribute Values in ECommerce Catalog by Learning with Limited Labeled Data.*
Yaqing Wang, Yifan Ethan Xu, Xian Li, Xin Luna Dong and Jing Gao.
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD 2020**).
Oral paper, acceptance rate = 5.8%

[14] *AutoKnow: Self-Driving Knowledge Collection for Products of Thousands of Types.*
Gabriel Blanco Saldana, Saurabh Deshpande, Xin Luna Dong, Xiang He, Andrey Kan, Xian Li, Yan Liang, Jun Ma, Alexandre Michetti Manduca, Jay Ren, Surender Pal Singh, Fan Xiao, Yifan Ethan Xu, Chenwei Zhang, Tong Zhao, Haw-Shiuan Chang, Giannis Karamanolakis, Yuning Mao, **Yaqing Wang**, Christos Faloutsos, Andrew McCallum and Jiawei Han.
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (**KDD 2020**).
Full paper, acceptance rate = 16.0%

[13] *Weak Supervision for Fake News Detection via Reinforcement Learning.*
Yaqing Wang, Weifeng Yang, Fenglong Ma, Jin Xu, Bin Zhong, Qiang Deng, Jing Gao.
Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (**AAAI 2020**).
Acceptance rate = 27.0%

[12] *Efficient Knowledge Graph Validation via Cross-Graph Representation Learning.*
Yaqing Wang, Fenglong Ma, Jing Gao.
Proceedings of ACM International Conference on Information and Knowledge Management (**CIKM 2020**).
Long paper, acceptance rate = 21.0%

[11] *Rare Disease Prediction by Generating Quality-Assured Electronic Health Records.*
Fenglong Ma*, **Yaqing Wang***, Jing Gao, Houping Xiao, Jing Zhou.
Proceedings of the SIAM International Conference on Data Mining (**SDM 2020**), Cincinnati, Ohio, May 7-9, 2020.
Acceptance rate = 24.0%

[10] *MeSHProbeNet: A Self-attentive Probe Net for MeSH Indexing.*
Guangxu Xun, Kishlay Jha, Ye Yuan, **Yaqing Wang** and Aidong Zhang.
Bioinformatics, Oxford University Press, 2019

[9] *Hypothesis Generation From Text Based On Co-Evolution Of Biomedical Concepts.*
Kishlay Jha, Guangxu Xun, **Yaqing Wang**, Aidong Zhang.
Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2019**).
Acceptance rate = 170/1200 = 14.2%

[8] *A General Framework for Diagnosis Prediction via Incorporating Medical Code Descriptions.*
Fenglong Ma, **Yaqing Wang**, Houping Xiao, Ye Yuan, Radha Chitta, Jing Zhou and Jing Gao.
Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (**BIBM 18**), Madrid, Spain, December, 2018.
Acceptance rate: 105/534 = 19.6%

[7] *Multivariate Sleep Stage Classification using Hybrid Self-Attentive Deep Learning Networks.*
Ye Yuan, Fenglong Ma, Guangxu Xun, **Yaqing Wang**, Kebin Jia, Lu Su and Aidong Zhang.
Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (**BIBM 18**), Madrid, Spain, December, 2018.
Acceptance rate: 105/534 = 19.6%

[6] *Interpretable Word Embeddings For Medical Domain.*
Kishlay Jha*, **Yaqing Wang***, Guangxu Xun, and Aidong Zhang.
Proceedings of the 18th IEEE International Conference on Data Mining (**ICDM 18**), Singapore, November 2018.
Acceptance rate: 19.94%

[5] *MuVAN: A Multi-view Attention Network for Clinical Temporal Data.*
Ye Yuan, Guangxu Xun, Fenglong Ma, **Yaqing Wang**, Nan Du, Kebin Jia, Lu Su and Aidong Zhang.
IEEE International Conference on Data Mining (ICDM 2018), Singapore, November 2018.
Acceptance rate: 8.86%

[4] *Towards Environment Independent Device Free Human Activity Recognition.*
Wenjun Jiang, Chenglin Miao, Fenglong Ma, Shuochao Yao, **Yaqing Wang**, Ye Yuan, Hongfei Xue, Chen Song, Xin Ma, Dimitrios Koutsonikolas, Wenyao Xu, and Lu Su.
The 24th Annual International Conference on Mobile Computing and Networking (**MobiCom 2018**), New Delhi, India, October 29-November 2, 2018.
Acceptance rate: 42/187 = 22%

[3] *EANN: Event Adversarial Neural Networks for Multi-Modal Fake News Detection.*
Yaqing Wang, Fenglong Ma, Zhiwei Jin, Ye Yuan, Guangxu Xun, Kishlay Jha, Lu Su and Jing Gao.
Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2018**), London, United Kingdom, August, 2018.
Acceptance rate: 112/496 = 22.6%

[2] *Concepts-Bridges: Uncovering Conceptual Bridges Based on Biomedical Concept Evolution.*
Kishlay Jha, Guangxu Xun, **Yaqing Wang**, Vishrawas Gopalakrishnan, Aidong Zhang.
Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (**KDD 2018**), London, United Kingdom, August, 2018.
Acceptance rate: 107/983 = 10.9%

[1] *Discovering Truths from Distributed Data.*
Yaqing Wang, Fenglong Ma, Lu Su, Jing Gao.
IEEE International Conference on Data Mining (**ICDM 2017**), New Orleans, USA, November 2017.
Acceptance rate: 72/778 = 9.25 %

HONORS & AWARDS

Bilsland Dissertation Fellowship, *Purdue University* 2022
Student Registration Award, *CIKM, Virtual* 2020
Student Registration Award, *KDD, Virtual* 2020
Student Travel Award, *AAAI, New York, USA* 2020
Student Travel Award, *KDD, London, UK* 2018
Student Travel Award, *ICDM, New Orleans, USA* 2017
Presidential Fellowship, *SUNY Buffalo* 2016-2020