

Wei (Victor) Yang

CONTACT INFORMATION	w85yang@uwaterloo.ca Homepage, Github, Linkedin, Google Scholar	
INDUSTRY EXPERIENCE	Research Engineer II Borealis.ai, Toronto, ON, Canada Machine Learning Engineer RSVP.ai, Waterloo, ON, Canada Software Engineer Intern Big Data Group, DiDi Research, Hangzhou, China Software Engineer Intern Pangu Studio, Netease Game, Hangzhou, China	Dec 2019 - Now May 2018 - Aug 2019 Feb 2017 - Jun 2017 June 2016 - Oct 2016
EDUCATION	University of Waterloo , Waterloo, ON, Canada M.Math., Computer Science (GPA: 86.8 / 100) <ul style="list-style-type: none">• Thesis: End-to-End Neural Information Retrieval [Paper] [Slide]• Thesis Supervisor: Jimmy Lin Singapore University of Technology and Design , Singapore Exchange Student (GPA: 4.0 / 4.0) Zhejiang University , Hangzhou, China B.S., Computer Science and Technology (GPA: 3.87 / 4.0) <ul style="list-style-type: none">• Thesis: Cross Domain Word Embedding• Thesis Supervisor: Lu Wei and Siliang Tang• Minor: Advanced Honors Class of Engineering Education (ACEE), Chu Kochen Honors College, Zhejiang University	Sep 2017 - Apr 2019 Sep 2016 - Dec 2016 Sep 2013 - Jun 2017
ACADEMIC EXPERIENCE	Research Assistant Data System Group, University of Waterloo, supervised by Jimmy Lin Research Areas: Neural Information Retrieval, Open-domain Question Answering Research Assistant Singapore University of Technology and Design, supervised by Lu Wei Research Areas: Cross-domain Word Embedding Research Assistant The Hong Kong Polytechnic University, supervised by Ajay Kumar (IEEE Fellow) Research Areas: Neural Palmprint Recognition	Sep 2017 - Apr 2019 Oct 2016 - Feb 2017 Jun 2016 - Sep 2016
AWARDS	<ul style="list-style-type: none">• Rank 1st in the MS-MARCO Passage Retrieval task• International Student Scholarship, University of Waterloo• Student Researcher Scholarship, EMNLP/SIGIR/NAACL• Graduate with Honors, Zhejiang University & Zhejiang Province• Champion, Zhejiang University Programming Contest, ACM-ICPC	2019 2018 2017/2019/2019 2017 2016
OPEN SOURCE CONTRIBUTION	Castor • Anserini • NCRFpp • Birch • MatchZoo • BERT-Fine_tune • Capreolus	
SERVICE	Sub-reviewer: ACL 2020, NAACL 2019, SIGIR 2019, CIKM 2019, AAAI 2018 Reviewer: AACL 2020, EMNLP(2020, 2019), COLING 2020, AAAI(2021, 2020), IJCAI 2021, NAACL 2021	

REFERRED PAPERS

1. Zhengkai Tu*, **Wei Yang***, Zihang Fu, Yuqing Xie, Luchen Tan, Kun Xiong, Ming Li and Jimmy Lin “Approximate Nearest Neighbor Search and Lightweight Dense Vector Reranking in Multi-Stage Retrieval Architectures” *ICTIR*, 2020. (* equal contribution) [[Paper](#)]
2. Rodrigo Nogueira, **Wei Yang**, Kyunghyun Cho, Jimmy Lin “Multi-stage document ranking with BERT” [[Paper](#)] [[Code](#)]
3. **Wei Yang***, Yuqing Xie*, Luchen Tan, Kun Xiong, Ming Li, and Jimmy Lin “Distant Supervision for Multi-Stage Fine-Tuning in Retrieval-Based Question Answering” (* equal contribution) *WWW*, 2020. [[Paper](#)]
4. Andrew Yates, Siddhant Arora, Xinyu Zhang, **Wei Yang**, Kevin Martin Jose and Jimmy Lin “Capreolus: A Toolkit for End-to-End Neural Ad Hoc Retrieval.” *WSDM demo*, 2020. [[Paper](#)] [[Code](#)]
5. Zeynep Akkalyoncu Yilmaz, Shengjin Wang, **Wei Yang**, Haotian Zhang and Jimmy Lin “Applying BERT to Document Retrieval with Birch.” *EMNLP demo*, 2019. [[Paper](#)] [[Code](#)]
6. Jinfeng Rao, Linqing Liu, Yi Tay, **Wei Yang**, Peng Shi, Jimmy Lin “Bridging the Gap of Relevance Matching and Semantic Matching with Hierarchical Co-Attention Network.” *EMNLP*, 2019. [[Paper](#)] [[Code](#)]
7. Linqing Liu, **Wei Yang**, Jinfeng Rao, Raphael Tang, and Jimmy Lin “Incorporating Compositional and Syntactic Structures Improves Semantic Similarity Modeling.” *EMNLP*, 2019. [[Paper](#)] [[Code](#)]
8. Zeynep Akkalyoncu, **Wei Yang**, Haotian Zhang, and Jimmy Lin “Cross-Domain Modeling of Sentence-Level Evidence for Document Retrieval.” *EMNLP*, 2019. [[Paper](#)] [[Code](#)]
9. Rodrigo Nogueira, **Wei Yang**, Jimmy Lin, and Kyunghyun Cho “Document Expansion by Query Prediction.” [[Paper](#)] [[Code](#)]
10. **Wei Yang***, Haotian Zhang*, and Jimmy Lin “Simple Applications of BERT for Ad Hoc Document Retrieval.” (* equal contribution) [[Paper](#)] [[Code](#)]
11. **Wei Yang**, Kuang Lu, Peilin Yang, Jimmy Lin “Critically Examining the “Neural Hype”: Weak Baselines and the Additivity of Effectiveness Gains from Neural Ranking Models.” *SIGIR*, 2019. (**Best Short Paper Honorable Mention**) [[Paper](#)] [[Code](#)]
12. **Wei Yang***, Yuqing Xie*, Aileen Lin, Xingyu li, Luchen Tan, Kun Xiong, Ming Li, and Jimmy Lin “End-to-End Open-Domain Question Answering with BERTserini.” (* equal contribution) *NAACL*, 2019. [[Paper](#)] [[Code](#)]
13. **Wei Yang**, Luchen Tan, Anqi Cui, Han Li, Xi Chen, Chunwei Lu, Muzi Wang, and Jimmy Lin “Detecting Customer Complaint Escalation with Recurrent Neural Networks and Manually-Engineered Features.” *NAACL industry track*, 2019. [[Paper](#)]
14. **Wei Yang**, Junnan Chen, Rui Qiao, Haocheng Qin, Ming Li, and Jimmy Lin “End-to-End Neural Context Reconstruction in Chinese Dialogue.” *ACL workshop (NLP for Conversational AI)*, 2019. [[Paper](#)]
15. Jinfeng Rao, **Wei Yang**, Yuhao Zhang, Ferhan Ture, and Jimmy Lin “Multi-Perspective Relevance Matching with Hierarchical ConvNets for Social Media Search.” *AAAI*, 2019. [[Paper](#)] [[Code](#)]
16. **Wei Yang**, Wei Lu, Vincent W. Zheng “A Simple Regularization-based Algorithm for Learning Cross-Domain Word Embeddings.” *EMNLP*, 2017. [[Paper](#)] [[Code](#)]