Wei (Victor) Yang

CONTACT INFORMATION	victor950118@gmail.com, w85yang@uwaterloo.ca Homepage, Github, Linkedin, Google Scholar

INDUSTRY EXPERIENCE	**Research Engineer II**	Dec 2019 to -
	Borealis.ai, Toronto, ON, Canada	
	Machine Learning Engineer	May 2018 to Aug 2019
	RSVP.ai, Waterloo, ON, Canada	
	Software Engineer Intern	Feb 2017 to Jun 2017
	Big Data Group, DiDi Research, Hangzhou, China	
	Software Engineer Intern	June 2016 to Oct 2016
	Pangu Studio, Netease Game, Hangzhou, China	

EDUCATION

University of Waterloo, Waterloo, ON, Canada

M.Math., Computer Science	Sep 2017 to Apr 2019

- Thesis: End-to-End Neural Information Retrieval [**Paper**] [**Slide**]
- Thesis Supervisor: Jimmy Lin

Singapore University of Technology and Design, Singapore

Exchange Student	Sep 2016 to Dec 2016

- GPA: 4.0 / 4.0 in two courses (Machine Learning, Optimization)

Zhejiang University, Hangzhou, China

B.S., Computer Science and Technology	Sep 2013 to Jun 2017

- Thesis: Cross Domain Word Embedding
- GPA: 3.87 / 4.0 (198.5 credits)
- Thesis Supervisor: Lu Wei and Siliang Tang

ACADEMIC EXPERIENCE	**Research Assistant**	Sep 2017 to Apr 2019
	Data System Group, University of Waterloo, supervised by Jimmy Lin	
	Research Assistant	Oct 2016 to Feb 2017
	Singapore University of Technology and Design, supervised by Lu Wei	
	Research Assistant	Jun 2016 to Sep 2016
	The Hong Kong Polytechnic University, supervised by Ajay Kumar (IEEE Fellow)	

AWARDS	• Top-one submission in the MS-MARCO Passage Retrieval task	2019
	• International Student Scholarship, University of Waterloo	2018
	• Graduate with Honors, Zhejiang University & Zhejiang Province	2017

OPEN SOURCE CONTRIBUTION	Castor • Anserini • NCRFpp • Birch • MatchZoo • BERT-Fine_tune • Capreolus

REFERRED PAPERS	1. Zhengkai Tu, **Wei Yang**, Zihang Fu, Yuqing Xie, Luchen Tan, Kun Xiong, Ming Li and Jimmy Lin "Content Analysis, Recommendation and Classification, Search and Ranking" *ICTIR*, 2020. 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 "Data Augmentation for BERT Fine-Tuning in Open-Domain Question Answering" (* equal contribution) *WWW*, 2020.

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. **[Code]**

5. Zeynep Akkalyoncu Yilmaz, Shengjin Wang, **Wei Yang**, Haotian Zhang and Jimmy Lin "Applying BERT to Document Retrieval with Birch." *EMNLP*, 2019.

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]**

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*, 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. **Wei Yang** "Pairwise Loss with Max Sampling: A Neural Meta-Architecture for Answer Selection." **[Code]**

16. 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]**

17. **Wei Yang**, Wei Lu, Vincent W. Zheng "A Simple Regularization-based Algorithm for Learning Cross-Domain Word Embeddings." *EMNLP*, 2017. **[Paper] [Code]**

18. Tingting Liang, Liang Chen, Zhining Xie, **Wei Yang**, Jian Wu "CASE: A Platform for Crowdsourcing Based API Search." *ICSOC*, 2015. **[Paper]**

SERVICE   Sub-reviewer: ACL 2020, NAACL 2019, SIGIR 2019, CIKM 2019, AAAI 2018
Reviewer: AACL 2020, EMNLP 2020, COLING 2020, AAAI 2020, EMNLP 2019