Title: Effective and Efficient Knowledge-Intensive NLP

Abstract:
Knowledge-intensive NLP tasks are the tasks that humans could not reasonably be expected to perform without access to external knowledge sources such as search engines, Wikipedia, dictionaries, and knowledge bases. They include open-domain question answering, commonsense reasoning, fact checking, etc. The state-of-the-art performance on such kinds of tasks is achieved by knowledge-augmented NLP solutions. They look for useful knowledge to augment the input for learning and prediction. However, the external data are heterogeneous and created independently from the task input; also, indexing and retrieval are expensive in time and space. In this talk, I will introduce our recent work in EMNLP 2022, ICLR 2023, and ACL 2023 on effective and efficient knowledge augmentation. It extends retrieval augmentation beyond unstructured data for language model training and usage. Since three conference tutorials in ACL/EMNLP and two successful workshops at AAAI 2023 and KDD 2023, this area of study has established a growing and enduring community. Join us if you are interested!

Bio:
Meng Jiang is an Associate Professor in the Department of Computer Science and Engineering at the University of Notre Dame. He received B.E. and PhD from Tsinghua University. He was a visiting PhD at CMU and a postdoc at UIUC. He is interested in data mining, machine learning, and natural language processing. His data science research focuses on graph and text data for applications such as question answering, query understanding, user modeling, material discovery, online education, and mental healthcare. He received the CAREER Award from the National Science Foundation. He has delivered 14 conference tutorials and organized 7 workshops. He is a Senior Member of ACM and IEEE.