2023 IEEE 5th International Conference on Cognitive Machine Intelligence (CogMI) | 979-8-3503-2383-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/COGMI58952.2023.00009

Keynote



Large Foundational Model Is a Blessing to Natural Language Understanding and Data Mining

Jiawei Han
University of Illinois at Urbana-Champaign, USA

Abstract

With the emergence of various large foundational models, it is important to examine their potential impact on natural language understanding and data mining. Although large language models (LLMs) may sometimes generate hallucinated contents, a proper usage of them will endow us with the power to uncovering hidden semantics in natural languages. We examine some recent studies on using large language models to enhance the power of natural language understanding and text mining, via a set of examples, including discriminative topic mining, text classification, and taxonomy-guided information extraction and construction of theme-specific knowledgebases. We show that equipped with LLMs, weakly supervised, annotation-free approach could be promising at transforming massive text into structured knowledge and benefiting many downstream applications. We will also discuss some unrealistic expectations and the issues on proper usage of large foundational models.

Bio

Jiawei Han is Michael Aiken Chair Professor in the Department of Computer Science, University of Illinois at Urbana-Champaign. He received ACM SIGKDD Innovation Award (2004), IEEE Computer Society Technical Achievement Award (2005), IEEE Computer Society W. Wallace McDowell Award (2009), Japan's Funai Achievement Award (2018), and was elevated to Fellow of Royal Society of Canada (2022). He is Fellow of ACM and Fellow of IEEE and served as the Director of Information Network Academic Research Center (INARC) (2009-2016) supported by the Network Science-Collaborative Technology Alliance (NS-CTA) program of U.S. Army Research Lab and co-Director of KnowEnG, a Center of Excellence in Big Data Computing (2014-2019), funded by NIH Big Data to Knowledge (BD2K) Initiative. Currently, he is serving on the executive committees of two NSF funded research centers: MMLI (Molecular Make Research Institute)—one of NSF funded national AI centers since 2020 and I-Guide—The National Science Foundation (NSF) Institute for Geospatial Understanding through an Integrative Discovery Environment (I-GUIDE) since 2021.