Yuhan Li

Shandong Province, China · (+86) 187-5439-9399 · yuhanli@mail.nankai.edu.cn · *Homepage*

Personal Information

My research interest is in knowledge graph and data mining. My current research focuses on **Entity Linking**. I have strong interests in knowledge graphs and their related applications, such as knowledge base question answering, knowledge-aware recommendation, and knowledge graph construction. Currently a third-year master candidate, studying at **Nankai University** under the supervision of the Associate Professor **Wei Shen**.

Education

Nankai University, Computer Science, *Master's Degree* 2020.09 - Present

(85.29 / 100), National Scholarship, GongNeng Scholarship

Northeast Forestry University, Software Engineering, *Bachelor's Degree* 2016.09 - 2020.06

(93.79 / 100), National Scholarship, Provincial-Level Merit Student (2 times), Outstanding Graduates

Research Experience

Entity Linking Meets Deep Learning: Techniques and Solutions (TKDE 2021),

Wei Shen*, **Yuhan Li**, Yinan Liu, Jiawei Han, Jianyong Wang, Xiaojie Yuan. 2020.02 - 2020.12

- Proposed a new taxonomy, which organizes more than fifty existing deep learning based entity linking models using three axes, i.e., embeddings, features, and algorithms.
- Discussed the remaining limitations of existing methods and highlighted some promising future directions.

Community Question Answering Entity Linking via Leveraging Auxiliary Data (IJCAI 2022),

Yuhan Li, Wei Shen*, Jianbo Gao, Yadong Wang. 2021.04 - 2022.01

- Explored the entity linking task based on CQA texts and constructed a finely-labeled data set named QuoraEL via crawling from Quora. Proposed a novel Transformer-based framework which can leverage different kinds of auxiliary data provided by CQA platforms effectively to enhance the linking performance.

Learning Entity Linking Features for Emerging Entities (TKDE 2022),

Chenwei Ran, Wei Shen*, Jianbo Gao, **Yuhan Li**, Jianyong Wang, Yantao Jia. 2021.01 - 2021.05

- Proposed a novel self-training based approach called STAMO to learn high-quality EL features for emerging entities automatically.

TIARA: Empowering Language Models on Question Answering over Large KBs (EMNLP 2022),

Yiheng Shu, Zhiwei Yu*, **Yuhan Li**, Börje Karlsson, Tingting Ma, Yuzhong Qu, Chin-Yew Lin 2022.04 - 2022.06

- Presented TIARA, which applys multi-grained retrieval to help the PLM focus on the most relevant KB contexts, viz., entities, exemplary logical forms, and schema items in KBQA task.

Professional Experience

MSRA, Knowledge Computing Group, Research Internship, Beijing 2022.04 - Present

- Co-advised by Zhiwei Yu and Börje Karlsson.
- From 2022.04 - 2022.06, researched in **Knowledge Base Question Answering**. From 2022.06 - present, researching in **Scientific Information Extraction** and **Scientific Knowledge Graph Construction**.

Sohu, Big Data Mining Group, Engineering Internship, Beijing 2019.08 - 2019.11

- Advised by Di Wang.
- Leveraged big data frameworks such as **Hadoop**, **Spark**, **Flink**, **Hive**, and **Kafka** to help develop a visual real-time advertising traffic monitoring system.

Skills & Others

- **Programming:** Python, Shell, C++, Java. Proficient in using PyTorch.
- **Language:** IELTS 6.5 (Reading 8.0, Writing 6.5, Listening 6.0, Speaking 6.0).
- **Leadership experience:** Served as the class monitor during the undergraduate period. Served as the Teaching Assistant of C++ course of Nankai University in 2020 Fall, 2021 Spring, and 2021 Fall.