RESEARCH STATEMENT

Ye Liu (liuyer@mail.ustc.edu.cn)

My research interests encompass a wide range of subjects within the fields of *Knowledge-aware Natural Language Processing (NLP)*, focusing on two main areas: 1) *Knowledge Acquisition* and 2) *Knowledge Application*, aiming to develop more powerful and reliable NLP systems. My approach emphasizes reducing the data dependence in the knowledge acquisition phase and focusing more on the generalization ability and robustness during the knowledge application phase. This results in a more versatile and transferable solution for natural language understanding and processing techniques. The overview of my research work is illustrated in Figure 1.

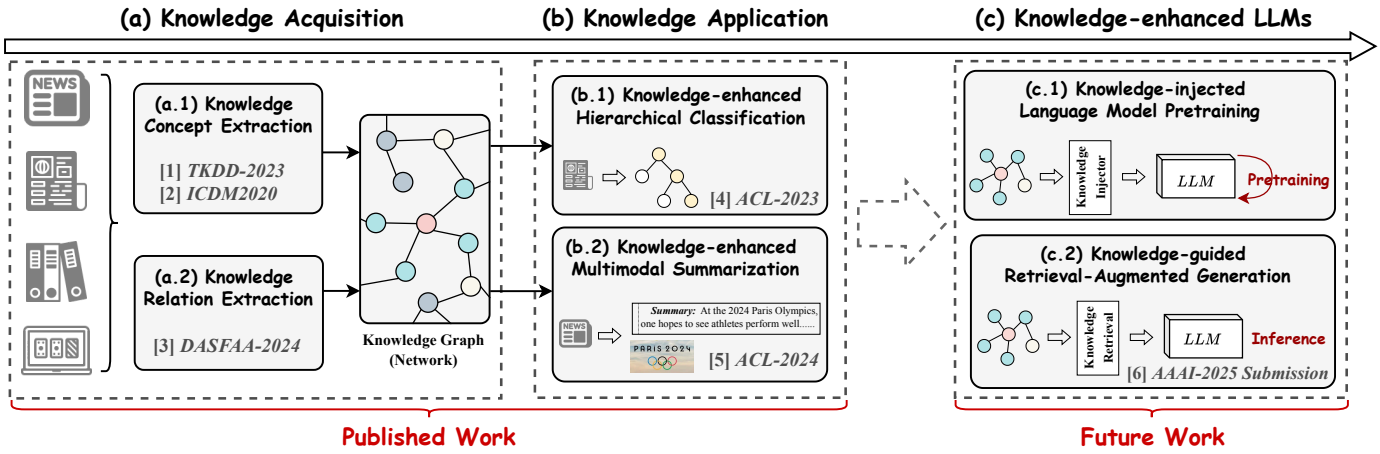

Figure 1: The Architecture of My Research

Knowledge Acquisition

My research aims to develop a series of solutions that require little to no labeled data to extract knowledge from unstructured documents. It focuses on two aspects including *Knowledge Concept Extraction* and *Knowledge Relation Extraction*.

1. Knowledge Concept Extraction

The first step of our research is to recognize the knowledge concepts contained in various unstructured documents. To achieve label-free and efficient knowledge concept recognition, I introduce a Hierarchical Multi-aspect Concept Extractor (HMCE) [1][2]. It takes account of the semantic characteristics of knowledge concepts and the potential structure of given documents, such as multi-aspect semantics and multi-level relevances. I further devise a graph-based concept ranking algorithm to assess the quality of candidate concepts, thus achieving robust and accurate knowledge concept recognition. This work has been published at *ICDM-2020* and *ACM TKDD-2023*.

2. Knowledge Relation Extraction

With the extracted knowledge concept, the next target is to determine the relation categories between different concepts. Traditional deep learning methods are significantly hampered by a lack of necessary prior knowledge, while large language models fall short in their task-specific capabilities for relation extraction. To address these shortcomings, I propose a Dual-System Augmented Relation Extractor (DSARE) [3], which synergistically combines traditional methods with LLMs. Specifically, DSARE imparts the prior knowledge inherent in LLMs to the traditional models while simultaneously transferring the traditional model's understanding of relation extraction to LLMs. In practice, DSARE demonstrates optimal performance, especially in low-resource settings. It was published at *DASFAA-2024*.

Knowledge Application

With the constructed knowledge network (knowledge graph), I am dedicated to exploring better interactions between external knowledge and downstream tasks, thereby making fuller use of the knowledge and achieving more competitive performance.

1. Knowledge-enhanced Hierarchical Classification Framework

Hierarchical Text Classification (HTC) is an essential and challenging subtask of multi-label text classification with a taxonomic hierarchy. Traditional methods have significant limitations due to the lack of domain knowledge, particularly for the classification at lower levels of the hierarchy. In my research [4], I attempt to incorporate the

extracted knowledge into the hierarchical classification process. Specifically, I innovatively integrate knowledge into both the text representation and hierarchical label learning process, addressing the knowledge limitations of traditional methods. Additionally, a novel knowledge-aware contrastive learning strategy is proposed to further exploit the information inherent in the data. This work has been published at *ACL-Findings-2023*.

2. **Knowledge-enhanced Multimodal Summarization**

   Multimodal Summarization with Multimodal Output (MSMO) aims to produce a multimodal summary with a textual abstract alongside a pertinent image. Traditional approaches typically adopt a holistic perspective on coarse image-text data, overlooking the essential connections between objects and the entities they represent. To address this, we propose an Entity-Guided Multimodal Summarization model (EGMS) [5]. EGMS utilizes dual multimodal encoders with shared weights to process text-image and entity-image information concurrently. Further, a gating mechanism combines visual data for knowledge entity-enhanced summary generation. Experiments validate its superiority and indicate the necessity of integrating external knowledge. This work has been published at *ACL-Findings-2024*.

Industry Application Experience

As a researcher with a passion for practical applications of NLP technology in the industry, I have had the honor of interning at Huawei Cloud & AI (Jan. 2019 – Aug. 2019) and ByteDance AI Lab (Feb. 2023 – Aug. 2023). At Huawei, I focused on developing an entity-guided question generation model. It employs the given entities in text as anchors to guide the question generation through a sequence-to-sequence (seq2seq) paradigm, overcoming the shortcomings of traditional question generation methods, such as irrelevant questions. At ByteDance, I explored the preference learning of LLMs, training a reward model with approximately 30,000 ranking pairs labeled by annotators. Then, using the RLHF (Reinforcement Learning from Human Feedback) algorithm, we optimized a finetuned generation model (13B size). This model has been deployed to generate summary descriptions of entities in the online encyclopedia knowledge graphs.

Future Research Direction

In future, I will continue exploring the knowledge application in various scenarios, with a focus on the Knowledge-enhanced Large Language Models, as illustrated in Figure 1 (c), mitigating the hallucination drawbacks of LLMs.

1. **Knowledge-injected Language Model Pretraining**

   Most existing LLMs, such as GPT-4 and Llama-3, are trained on various text corpora, including books and documents. This paradigm relies on LLMs to learn knowledge from unstructured documents and store it in the model parameters. However, due to the capacity limitations of LLMs and the forgetting phenomenon during the learning process, these LLMs often fail when faced with many domain-specific cases and situations requiring complex reasoning abilities. I will attempt to inject structured knowledge, including but not limited to domain knowledge triples, commonsense, long-tail and long-context knowledge, into the pretraining process of LLMs. This approach aims to enhance the complex reasoning abilities of LLMs, resulting in more powerful and reliable models.

2. **Knowledge-guided Retrieval-Augmented Generation (RAG)**

   The training process of LLMs requires powerful computing sources, which are often beyond the reach of most researchers worldwide. Fortunately, the powerful reasoning abilities of LLMs offer another way to integrate external knowledge: Retrieval-Augmented Generation (RAG). I have already made attempts in this direction, such as fake news detection [6]. Along this line, a key issue is how to ensure the correctness of retrieved information, especially in this age of disinformation. I will propose a knowledge-guided RAG system, using the concrete knowledge from knowledge graphs to judge and select retrieved information and address conflicts between different sources.

References

[1] Ye Liu, et al. TechPat: Technical Phrase Extraction for Patent Mining, *ACM TKDD, 2023*.

[2] Ye Liu, et al. Technical Phrase Extraction for Patent Mining: A Multi-level Approach, *ICDM, 2020*.

[3] Ye Liu, et al. Empowering Few-Shot Relation Extraction with Integration of Traditional Methods and LLMs, *DASFAA, 2024*.

[4] Ye Liu, et al. Enhancing Hierarchical Text Classification through Knowledge Graph Integration, *ACL-Findings, 2023*.

[5] Yanghai Zhang, Ye Liu, et al. Leveraging Entity Information for Cross-Modality Correlation Learning: The Entity-Guided Multimodal Summarization, *ACL-Findings, 2024*.

[6] Ye Liu, et al. Detect, Investigate, Judge and Determine: A Novel LLM-based Framework for Few-shot Fake News Detection, Submitted to *AAAI, 2025.*