Shengkun Ma

No 10, Xitucheng Road, Haidian District, Beijing, China 100876
+86 1500-2972-032 mashengkun98@gmail.com mashengkun.github.io

Education

Beijing University of Posts and Telecommunications **Sep. 2021 – Jun. 2024**
M.S. in Computer Science and and Technology, School of Computer Science *Beijing, China*
Lab: State Key Laboratory of Networking and Switching Technology, supervised by Prof. Bo Cheng
Honor: *National Scholarship, Outstanding Graduate Student*

Xidian University **Sep. 2017 – Jun. 2021**
B.S. in Software Engineering, School of Computer Science and Technology *Xian, China*
Honor: *Second-class Scholarship with GPA 3.7/4.0*

Publications and Arxiv

- **MRCEval: A Comprehensive, Challenging and Accessible Machine Reading Comprehension Benchmark**
  Shengkun Ma, Hao Peng, Lei Hou and Juanzi Li.
  Under review
- **Making Pre-trained Language Models Better Continual Few-shot Relation Extractors**
  Shengkun Ma, Jiale Han, Yi Liang and Bo Cheng.
  Accepted as COLING 2024
- **A Rationale-Centric Data Augmentation Method for Cross-Document Event Coreference**
  Bowen Ding*, Qingkai Min*, **Shengkun Ma**, Linyi Yang and Yue Zhang.
  Accepted as NAACL 2024
- **Generative Prompt Tuning for Relation Classification**
  Jiale Han, Shuai Zhao, Bo Cheng, **Shengkun Ma** and Wei Lu.
  EMNLP Findings 2022

Research

Machine Reading Comprehension. **Jul. 2024 – Feb. 2025**
First author (Supervised by Prof. Juanzi Li and Dr. Hao Peng) *Beijing, China*

- This work summarizes machine reading comprehension datasets, proposes a novel MRC taxonomy and constructs a comprehensive, challenging and accessible MRC benchmark.
- We perform an extensive evaluation of LLMs, highlighting that MRC continues to present significant challenges even in the era of LLMs.
- This work was completed during my internship at KGE lab, Tsinghua University.
- *Under review.*

Making PLMs Better Continual Few-shot Relation Extractors. **Jan. 2023 – Sep. 2023**
First author (Supervised by Prof. Bo Cheng) *Beijing, China*

- This work focuses on continual few-shot relation extraction with two challenges: catatrophic forgetting and overfitting, and propose a contrastive prompt learning framework and introduce a memory augmentation strategy based on ChatGPT.
- We outperform current SOTA by a large merge and significantly mitigates catastrophic forgetting and overfitting.
- I'm the main person in charge of this project and this is my first work.
- *Accepted as COLING 2024.*

A Rationale-centric Data Augmentation Method for Cross-Document ECR. **May 2023 – Aug. 2023**
Group member of WestlakeNLP (Supervised by Prof. Yue Zhang and Dr. Linyi Yang) *Hangzhou, China*

- To reduce model overfitting on lexical surface matching of event triggers, we design a rationale-centric data augmentation method by generating trigger-diverse counterfactually augmented data from LLMs.
- I'm responsible for evaluation of LLMs for Cross-Document Event Coreference Resolution.
- *Accepted as NAACL 2024.*

Generative Prompt Tuning for Relation Classification. **Sep. 2021 – Jan. 2022**
Co-author with Dr.Han (Supervised by Prof. Bo Cheng) *Beijing, China*

- To apply prompt learning to relation classification, we propose a novel generative prompt tuning method to reformulate relation classification as an infilling problem and design an entity-guided decoding for inference.
- I'm mainly responsible for coding and conducting experiments.
- *EMNLP 2022 Findings.*

Experience

Benchmark of Machine Reading Comprehension **Jul. 2024 – Present**

Intern in KGE lab, Tsinghua University (Supervised by Prof. Juanzi Li and Dr. Hao Peng) *Beijing, China*

* Concentrating on the evaluation of reading comprehension ability for large language models. We summarize the current MRC datasets, propose a new MRC taxonomy, and develop a comprehensive MRC benchmark.
* I cover this project and work with Dr. Hao Peng.

Evaluation of Large Language Models **Apr. 2023 – Oct. 2023**

Intern in WestlakeNLP, Westlake University (Supervised by Prof. Yue Zhang and Dr. Linyi Yang) *Hangzhou, China*

* Participate in the evaluation of large language models and responsible for designing schemes to solve some downstream tasks with LLMs and evaluating the performance of LLMs.
* Some research work about Explainable Artificial Intelligence (XAI) and commonsense reasoning.

Skills Summary

Programming Languages: Python
Developer Tools: Linux, Pytorch, Huggingface
Languages: Mandarin, English