Shiping Yang

Actively seeking a PhD position for the 2024 fall

yangshiping@bupt.edu.cn Shiping Yang's linkedin https://maybenotime.github.io

EDUCATION

Beijing University of Posts and Telecommunications **Beijing, China**
Master of Engineering in Computer technology, GPA:3.81/4.0 *09/2022 - Present*

Beijing Language and Culture University **Beijing, China**
Bachelor of Engineering in Computer Science, GPA:3.36/4.0 *09/2018 - 06/2022*

- **Awards:**
  The 1st Prize winner (provincial level) in "Contemporary Undergraduate Mathematical Contest in Modeling"
  The Second Prize winner (national level) in "Chinese Undergraduate Computer Design Contest"

PUBLICATION

A New Benchmark and Reverse Validation Method for Passage-level Hallucination Detection **EMNLP2023**
Shiping Yang, *Renliang Sun, Xiaojun Wan* *Findings*

Multi-level Contrastive Learning for Scripts-based Character Understanding **EMNLP2023**
Dawei Li, Hengyuan Zhang, Yanran Li, ***Shiping Yang*** *Main*

A New Dataset and Empirical Study for Sentence Simplification in Chinese **ACL2023**
Shiping Yang, *Renliang Sun, Xiaojun Wan* *Main*

Fine-grained Contrastive Learning for Definition Generation **AACL2022**
Hengyuan Zhang, Dawei Li, ***Shiping Yang***, *Yanran Li* *Oral*

EXPERIENCE

Peking University **Beijing**
Research Intern *10/2022 - 08/2023*

- Conducted research on **Text Simplification**
- Conducted research on **Hallucination of Large Language Models**

Algorithm R&D Department of Tomorrow Advancing Life **Beijing**
Natural Language Processing Engineering Intern *06/2022 - 10/2022*

- Worked on **Reverse Dictionary** in Chinese and **deployed it online** to serve customers
- Participated in the "The fourth NiuTrans Cup humor computing" competition and won the **First Prize**

Beijing Advanced Innovation Center for Language Resources **Beijing**
Research Intern *11/2021 - 05/2022*

- Conducted research on **Hard-Constrained Text Generation**
- Developed a website using **Vue** and **Tornado** to demonstrate the effect of models

RESEARCH

A New Benchmark and Reverse Validation Method for Passage-level Hallucination Detection **PDF**

- We design a two-stage annotation process and **entity select strategy** to create a high-quality and **challenging benchmark** named PHD for the evaluation of **passage-level hallucination detection**.
- We propose a **self-check method** based on reverse validation to detect passage-level hallucinations, which can be used in **black-box models** and **zero-resource fashion**.
- We implement **two variants** of our method on ChatGPT and **Llama-2-7b-chat-hf**, evaluating them on two datasets. The experimental results demonstrate that the proposed method outperforms existing methods by **a large margin**.

Multi-level Contrastive Learning for Scripts-based Character Understanding **PDF**

- We highlight the challenge of **script-based character understanding** from the perspective of **text length** and **text type** and propose a **multi-level contrastive learning framework** to address them.
- We introduce **summary-conversation contrastive learning** to help models understand the fine-grained information at the **in-sample level**, and design **cross-sample contrastive learning** to capture the global dependency at the **cross-sample level**.

- We empirically evaluate the proposed method on three character understanding sub-tasks. The experimental results demonstrate that our method is **effective** and **compatible** with previous SOTA models.

A New Dataset and Empirical Study for Sentence Simplification in Chinese **PDF**

- We create **a high-quality dataset** named CSS for the evaluation of Chinese SS models to facilitate research on Chinese sentence simplification.
- We conduct **data analysis** to compare the characteristics of CSS with English datasets, pointing out the difference between Chinese and English sentence simplification tasks.
- We exploit **transfer learning** and **unsupervised learning** to train Chinese SS models in low-resource scenarios and explore whether **LLMs** can serve as high-quality Chinese sentence simplification systems by evaluating them on CSS.

Fine-grained Contrastive Learning for Definition Generation **PDF**

- We point out the **under-specific problem** in the definitions generated by the current SOTA pre-trained encoder-decoder models and analyzed the reasons for that phenomenon.
- We propose a novel **fine-grained contrastive learning method** to align the representation of the word and definitions, encouraging the model to capture more detailed semantic components of the target word.
- We apply our method to a **T5 backbone** and conduct experiments on three popular datasets. The results demonstrate that the proposed method could generate more specific and high-quality definitions compared with several SOTA models.

PROJECTS

Reverse Dictionary in Chinese **07/2022 - 10/2022**

- We explore both **classification and generation methods** for the reverse dictionary task. Furthermore, we construct a private validation dataset using company data to choose models.
- We replicate the method proposed by ≪BERT for Monolingual and Cross-Lingual Reverse Dictionary≫ and improve its performance by **injecting external knowledge**.
- We **retrieve similar samples** from training data to augment user input, **bridging the gap** between user input and training data.

CCL2022 Evaluation Task — The fourth NiuTrans Cup humor computing **07/2022 - 09/2022**

- This evaluation task is divided into two tracks: **humor recognition** and **humor response generation** in dialogue scenarios. We rank top at both tracks and win the **First Prize** finally.
- For humor recognition, we explore sentence-based text classification, **dialogue modeling**, and **prompt-based** text classification. We conduct experiments using various **BERT variants** and apply a range of **tricks**. Then, we ensemble heterogeneous models through a **stacking** approach to further improve the performance.
- For humor response generation, we design a **sliding window strategy** to perform **data augmentation** and conduct experiments on GPT2 and T5. Furthermore, we utilize the humor recognition model to **select humorous response** from sampled results.

Hard-Constrained Text Generation with Controllable Word Complexity (Thesis) **03/2022 - 05/2022**

- We introduce **hard-constrained text generation** and **style transfer** to the field of international Chinese language education, generating example sentences based on user-entered keywords with **controllable word complexity**.
- We employ the T5 model to generate example sentences in a **fill-in-the-blank** way, ensuring that the generated sentences **include all the keywords**.
- We propose an **unsupervised method** to mine training data automatically from corpus. And we train the model using **prompt tuning** instead of standard fine-tuning to achieve style transfer at a **lower cost**.

SERVICES

Reviewer: Serve as a Reviewer for **ACL ARR**

Secondary Reviewer: Serve as a Secondary Reviewer for **ACL2023** and **EMNLP2023**

PRESENTATION

EMNLP2023: Attend **EMNLP2023** to present my work ***A New Benchmark and Reverse Validation Method for Passage-level Hallucination Detection*** in person as a **poster** format