JUNHAO LIU

Email: jhliu17@foxmail.com ⋄ Tel: (+86) 135 6003 7579

EDUCATION

University of Chinese Acadamy of Sciences (UCAS) Sep. 2019 - Jun. 2022

M.Eng. in Computer Science & Technology
Affiliation: SIAT, CAS · Adavisor: Prof. Min Yang
Natural Language Processing (Cross-lingual, Question Answering, Multi-modal, etc.)

South China University of Technology (SCUT) Sep. 2015 - Jun. 2019

B.Eng. in Intelligence Science & Technology
GPA: 3.79/4.0 · Rank: 2/51
Focus on Pattern Recognition, Machine Learning, and Robotics

EXPERIENCE

DAMO Academy, Alibaba Sep. 2020 - Feb. 2021

Research Intern · Adavisor: Lidong Bing

Bing NLP, Microsoft Mar. 2020 - Jul. 2020

Research Intern · Adavisor: Linjun Shou

SIAT, Chinese Academy of Sciences Mar. 2019 - Sep. 2019

Research Intern · Adavisor: Prof. Min Yang

PUBLICATIONS

- **Junhao Liu**, Linjun Shou, Jian Pei, Ming Gong, Min Yang and Daxin Jiang. Cross-lingual Machine Reading Comprehension with Language Branch Knowledge Distillation. In *Proceedings of the 28th International Conference on Computational Linguistics* (COLING-2020).
- Jian Wang, **Junhao Liu**, Wei Bi, Xiaojiang Liu, Ruifeng Xu, Kejing He and Min Yang. Dual Dynamic Memory Network for End-to-End Multi-turn Task-oriented Dialog Systems. In *Proceedings of the 28th International Conference on Computational Linguistics* (COLING-2020).
- **Junhao Liu**, Kai Wang, Chunpu Xu, Zhou Zhao, Ruifeng Xu, Ying Shen and Min Yang. Interactive Dual Generative Adversarial Networks for Image Captioning. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence* (AAAI-2020).
- Jian Wang, **Junhao Liu**, Wei Bi, Xiaojiang Liu, Kejing He, Ruifeng Xu and Min Yang. Improving Knowledge-aware Dialogue Generation via Knowledge Base Question Answering. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence* (AAAI-2020).
- **Junhao Liu**, Min Yang, Chengming Li, Ruifeng Xu. Improving Cross-Modal Image-Text Retrieval with Teacher-Student Learning. *IEEE Transactions on Circuits and Systems for Video Technology* (TCSVT, JCR Q1).

RESEARCH PROJECTS

Cross-lingual Machine Reading Comprehension Mar. 2020 - Jul. 2020

Research project at Bing NLP Group, Microsoft

- Propose an augmentation approach named Language Branch Machine Reading Comprehension (LBMRC) by training several language-specific models as teachers to provide fine-grained supervisions. With the LBMRC, we can train multiple MRC models proficient in individual language based on translation data. Then, we devise a knowledge distillation approach to transfer knowledge from separate language branch models to a unified Cross-lingual Machine Reading Comprehension (CLMRC) model for all target languages, to save the cost of training, inference, and maintenance for multiple models.

· Extensive experiments on two popular CLMRC benchmark datasets in 9 languages under both translation and zero-shot conditions. Our model achieves state-of-the-art results on all languages for both datasets without using any external large-scale corpus which verifies the effectiveness of LBMRC and multilingual distillation framework.

End-to-end Task-oriented Dialog Systems May. 2019 - Sep. 2019

Research project at SIAT, Chinese Academy of Sciences

· A Dual Dynamic Memory Network (DDMN) framework for knowledge reasoning in multi-turn task-oriented dialog systems, which dynamically keeps track of long-term dialog states and KB knowledge with a novel updating mechanism. In this project, we propose to utilize the Self-critical Sequence Training (SCST) strategy to boost the KB entity selection performance of the decoder policy.
· Experimental results on In-Car Assistant and CamRest datasets show that the proposed model achieves impressive results compared to the existing methods across multiple evaluation metrics. We also perform a detailed analysis to study DDMN, showing that the model is able to maintain more sustained conversations, and performs much better than the compared methods with the increase of dialog turns.

Image Captioning with Dual Interaction Mar. 2019 – Sep. 2019

Research project at SIAT, Chinese Academy of Sciences

· Propose an Interactive Dual Generative Adversarial Network (IDGAN) for image captioning, which mutually combines the retrieval-based and generation-based methods to learn a better image captioning ensemble. IDGAN consists of two generators and two discriminators, where the generation- and retrieval-based generators mutually benefit from each other's complementary targets that are learned from two dual adversarial discriminators.
· Extensive experiments on MSCOCO dataset demonstrate that the proposed IDGAN model significantly outperforms the compared methods for image captioning.

Abstractive Text Summarization Aug. 2018 - Dec. 2018

Competition project of 2018 Byte Cup International Machine Learning Competition

· Worked in a team to solute the task of automatically generating titles for given articles in Byte Cup 2018 International Machine Learning Competition, where we proposed to employ a CNN network to select salient sentences, then employ a sequence-to-sequence model with copy mechanism to rewrite the extracted sentences, finally build an advantage actor-critic (A2C) model with policy gradient to improve the performance.
· Achieved Rouge-L score of 35.38 on online validation and 38.98 on online test, awarded as the 3rd prize in the finals among over 1000 teams.

HONORS/AWARDS

AAAI 2020 Student Scholarship 2020
3rd Prize, Award on Byte Cup 2018 International Machine Learning Contest Dec. 2018
SCUT Scholarship 2016, 2018
Honor Prize, Award on 2018 Mathematical Contest In Modeling 2018
Endress + Hauser SC China Scholarship 2017
2nd Prize, Award on 2017 China Undergraduate Mathematical Contest in Modeling 2017
Merit Student of SCUT Award & Outstanding Student Leader Award 2016, 2018

SERVICES

Reviewer COLING 2020, IEEE Transactions on Cognitive and Developmental Systems

TECHNICAL SKILLS

Programming Languages Python, C/C++, Java, MATLAB, Shell
Tools PyTorch, TensorFlow, Keras, Git, Transformers
Others LaTeX, Adobe After Effect and Premiere, Microsoft Office