

# Yawei Sun

13382055988 | [ywsun@ir.hit.edu.cn](mailto:ywsun@ir.hit.edu.cn) | [Yawei Sun](#) | <https://github.com/syw1996>

## EDUCATION

### Harbin Institute of Technology

*Master of Computer Science and Technology, Minor in NLP*

Harbin, China

*Sep. 2019 – Jul 2021*

### Harbin Institute of Technology

*Bachelor of Computer Science and Technology*

Harbin, China

*Sep. 2015 – Jul 2019*

## EXPERIENCE

### Research Intern

*Pengcheng Lab*

Feb 2019 – May 2019

*Shenzhen, China*

- Pengcheng Lab is currently a second-class institution approved by the Guangdong Provincial Government and built by the Shenzhen Municipal Government
- Took part in the research on Table-to-Text generation based on the Rotowire dataset and the work's contribution has been written in the paper named **Learning to Select Bi-Aspect Information for Document-Scale Text Content Manipulation** which has been accepted by AAAI 2020

### NLP Engineer Assistant

*Harbin DeepIntell Company Limited*

Aug 2018 – Dec 2018

*Harbin, China*

- Implemented an retrieval automatic comment system based on paper **Multiway Attention Networks for Modeling Sentence Pairs**, github link: <https://github.com/syw1996/Retrieval-Automatic-Comment-System>
- Implemented a keyword chatbot based on paper **Sequence to Backward and Forward Sequences: A Content-Introducing Approach to Generative Short-Text Conversation**, github link: <https://github.com/syw1996/Seq2BF-pytorch>
- Migrated code from TensorFlow to Pytorch, based on paper **Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory**

## PUBLICATIONS

### GPT-based Few-shot Table-to-Text Generation with Table Structure Reconstruction and Content Matching

- Heng Gong\*, **Yawei Sun\***, Xiaocheng Feng, Bing Qin, Wei BI, Xiaojiang Liu, Ting Liu. Proceedings of COLING 2020 (Long)

### Learning to Select Bi-Aspect Information for Document-Scale Text Content Manipulation

- Xiaocheng Feng, **Yawei Sun**, Bing Qin, Heng Gong, Yibo Sun, Wei BI, Xiaojiang Liu, Ting Liu. Proceedings of AAAI 2020 (Long)

## PROJECTS

### Tencent Novel Generation | *Pytorch, OpenNMT*

Sep 2018 – Feb 2019

- Designed an end-to-end novel generation model which use the appearance keywords to generate the novel character description fragment
- Introduced policy gradient which is a reinforcement learning mechanism and use the text bleu score as the reward
- Tried an unsupervised style transfer mechanism which is firsted proposed by Zhiting Hu and had a great improvement in the richness of the style of novel generation
- Applied to text generation platform which won innovation group second prize in the 4th National Youth Artificial Intelligence Innovation and Entrepreneurship Conference
- Successfully applied for Tencent's patent "An invention of paragraph-level automatic appearance description generation based on deep learning technolog"

## TECHNICAL SKILLS

**Languages:** Python (Proficient), C/C++ (Basic), Java (Basic), JavaScript (Basic)

**Frameworks:** Pytorch (Proficient), Tensorflow (Basic), Flask (Basic)

**Developer Tools:** Git, Conda, VS Code, PyCharm

**Languages:** English (CET-6 550, TOEFL 84)