



General Information

Full Name Haoyuan Peng - 彭浩源
Date of Birth 9th July 1994
Email phy_fdu@163.com
Homepage haoyuanpeng.github.io

Education

2015 - 2018

Master's degree, Computer Science

Software School of Fudan University

- Under the supervision of Prof. Zheng Xiaoqing, I studied state-of-the-art NLP research. Our research interests mainly lie in pre-training word embeddings and dependency parsing. Our works are published in AAAI-2017 and AAAI-2018.
- I participated in the development of FudanDNN-NLP, a deep learning-based natural language processing tool.
- I was awarded a postgraduate freshman scholarship in 2015.
- I was selected as an Outstanding Graduate in Shanghai in 2018.

2011 - 2015

Bachelor's degree, Computer Software Engineering

Software School of Fudan University

- Under the supervision of Prof. Zheng Xiaoqing, I have been conducting research on natural language processing since 2013. Our work is published in IJCAI-2015.
- I was awarded the Third Prize of the Scholarship for Outstanding Students at Fudan University in the 2014-2015 academic year.
- I have participated in Fudan's Undergraduate Experiment and Practice Opportunities Program from 2013 to 2015 under the guidance of Prof. Zheng Xiaoqing, and passed the thesis defence.

Experience

2023 - PRESENT

Senior Algorithm Engineer

Learnable.AI, Shanghai, China

- Engaging in applied research on the direction of K-12 auto-grading system. Transforming large language models (LLMs) and seq2seq models into real-world applications, such as translating arbitrary answers written by students in natural language into an internally defined formal language.

2021 - 2022

Senior Researcher

Tencent, Shanghai, China

- **Yunzhi Media AI Platform** - Responsible for the text error correction, video labeling and key information extraction algorithms of multimodal content structuring in Tencent Yunzhi Media AI Platform, which has won the *CCBN 2021 Product Innovation Excellence Award*. The labeling algorithm I was responsible for won the 2nd place in the *2021 AIWIN Algorithm Technology Competition* without training with the competition training data.
- **Text Censorship** - Developing a text censorship system for private project clients. The system contains models and policies, and can address adversarial user inputs.

2018 - 2021

Applied Researcher

Tencent, Shanghai, China

- **Intelligent Public Opinion Analysis System** - Responsible for the overall algorithm scheme and completed the development and optimization of multiple NLP algorithms. This product is customized for securities industry regulators.
- **Tencent Cloud NLP Products** - Responsible for the development of general text matching and Chinese spelling correction algorithms, which are deployed on Tencent Cloud and applied in several custom projects.
- **Tencent TI-ONE ML Platform** - Implemented multiple traditional ML algorithms and deep learning based NLP algorithms on Tencent TI-ONE ML Platform for users to train models on their custom data.

2014 - 2015

Data Analyst Intern

eBay, Shanghai, China

Skills

Following the cutting-edge research in the field of natural language processing. Experienced in publishing papers in top international academic conferences.

Knowledge in full life cycle AI/ML with learning, inference, engineering and integration.

Proficient in Python and able to write high-quality Python code.

Familiar with mainstream deep learning frameworks, including pytorch, tensorflow 1.x and tensorflow 2.x.

Service

Conference Reviewer: KDD-2023

External Reviewer: ACL-2023, ECAI-2023

Honors and Awards

2021

- Second Place of AIWIN 2021 Algorithm Technology Competition

2020

- Tencent New Code Culture Award - Award for outstanding internal open source code projects

2018

- Outstanding Graduate in Shanghai

2015

- Postgraduate Freshman Scholarship

Publications

2023

OSAN: A One-Stage Alignment Network To Unify Multimodal Alignment and Unsupervised Domain Adaptation

Liu, Ye and Qiao, Lingfeng and Lu, Changchong and Yin, Di and Lin, Chen and Peng, Haoyuan and Ren, Bo. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

2022

Grafting Pre-trained Models for Multimodal Headline Generation

Qiao, Lingfeng and Wu, Chen and Liu, Ye and Peng, Haoyuan and Yin, Di and Ren, Bo. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: Industry Track*.

2019

Detecting Abnormal Start-Ups, Unusual Resource Consumptions of the Smart Phone: A Deep Learning Approach

ZHENG, Xiaoqing and LU, Yaping and PENG, Haoyuan and FENG, Jiangtao and ZHOU, Yi and JIANG, Min and MA, Li and ZHANG, Ji and JI, Jie. *ZTE Communications*.

2018

Attention-based belief or disbelief feature extraction for dependency parsing

Peng, Haoyuan and Liu, Lu and Zhou, Yi and Zhou, Junying and Zheng, Xiaoqing. *Proceedings of the AAAI Conference on Artificial Intelligence*.

2018

RNN-based sequence-preserved attention for dependency parsing

Zhou, Yi and Zhou, Junying and Liu, Lu and Feng, Jiangtao and Peng, Haoyuan and Zheng, Xiaoqing. *Proceedings of the AAAI Conference on Artificial Intelligence*.

2017

Learning context-specific word/character embeddings

Zheng, Xiaoqing and Feng, Jiangtao and Chen, Yi and Peng, Haoyuan and Zhang, Wenqing. *Proceedings of the AAAI Conference on Artificial Intelligence*.

2015

Character-based parsing with convolutional neural network

Zheng, Xiaoqing and Peng, Haoyuan and Chen, Yi and Zhang, Pengjing and Zhang, Wenqiang. *Twenty-Fourth International Joint Conference on Artificial Intelligence*.