



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

- Conducted cutting-edge research in NLP under the mentorship of Professor Zheng Xiaoqing, with a focus on pre-training word embeddings and dependency parsing. Our work resulted in publications at conferences including AAI-2017 and AAI-2018.
- Played a significant role in the development of FudanDNN-NLP, a deep learning-based tool for NLP applications.
- Recognized as an Outstanding Graduate in Shanghai for the year 2018.

2011 - 2015

Bachelor's degree, Computer Software Engineering

Software School of Fudan University

- Engaged in NLP research under the guidance of Professor Zheng Xiaoqing since 2013. Our work resulted in a publication at IJCAI-2015.
- Successfully completed the Fudan Undergraduate Experiment and Practice Opportunities Program from 2013 to 2015, working closely with Professor Zheng Xiaoqing and passing the thesis defense.

Experience

2023 - PRESENT

Senior Algorithm Engineer

Learnable.AI, Shanghai, China

- Conduct research on enhancing the reasoning error detection capabilities of LLMs through Chain-of-Thought (CoT) technology. Our work has been submitted to IJCAI 2024.
- Train large-scale models ranging from 7B to 70B parameters for real-world systems in the education domain. Applications include directly grading students' mathematical free-response answers and translating student responses into internally defined languages.
- Investigate OCR result correction algorithms for scenarios involving student responses, effectively addressing the challenge of distinguishing between student writing errors and OCR recognition errors.
- Evaluated as Excellent in the performance evaluation of probation period.

2021 - 2022

Senior Researcher

Tencent, Shanghai, China

- Led the development of multiple video structuring algorithms within the **Yunzhi Media AI Platform**, encompassing multimodal key information extraction, video tagging, as well as ASR and OCR result correction. The tagging algorithm achieved 2nd place in the 2021 AIWIN Algorithm Technology Competition without utilizing the competition training data.
- Developed a **text moderation system** for private project clients, comprising models and policies capable of addressing adversarial user inputs.
- Recognized as an 'Outstanding Contributor' (5-star performance) in performance evaluation.

2018 - 2021

Applied Researcher

Tencent, Shanghai, China

- Implemented traditional ML algorithms and deep learning-based NLP algorithms on the **Tencent TI-ONE ML Platform**, enabling users to train models on their custom data.
- Contributed to the development of text matching and Chinese spelling correction algorithms deployed as **Tencent Cloud NLP Products** and applied in various custom projects.
- Spearheaded the development and optimization of multiple NLP algorithms as part of the **Public Opinion Analysis System**, tailored for securities industry regulators.
- Received a 4-star performance rating in performance evaluations on 2 occasions.

2014 - 2015

Data Analyst Intern

eBay, Shanghai, China

Skills

Proficient in conducting cutting-edge research in natural language processing, with a track record of publishing papers in top conferences.

Experienced in all aspects of the full life cycle of AI/ML projects, including training, inference, engineering, and integration.

Highly skilled in Python programming, capable of writing high-quality python codes.

Familiar with mainstream deep learning frameworks, including PyTorch, TensorFlow 1.x, and TensorFlow 2.x.

Service

Conference Reviewer: KDD-2023, EMNLP-2023, SDM-2024, COLING-2024

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

Publications

2023

VKIE: The Application of Key Information Extraction on Video Text

An, Siyu and Liu, Ye and Peng, Haoyuan and Yin, Di. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track*.

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*.