

# Pai (Patrick) Liu

Westlake University, Hangzhou, China  
(086)189-7119-8121 | [liupai@westlake.edu.cn](mailto:liupai@westlake.edu.cn)

## EDUCATION

### Northeastern University

Shenyang, China

**Bachelor** of Software Engineering, GPA 3.86/4.0, grade 90.7/100, top 3%

Sep 2017 – Jun 2021

**Core Courses:** Natural Language Processing(4.0/4.0), Recommendation System(4.0/4.0), Linear Algebra(4.0/4.0), Probability and Statistics(4.0/4.0), Neural Networks and Deep Learning(cousera), Improving Deep Neural Networks(cousera)

### Westlake University (Visiting Student)

Hangzhou, China

School of Engineering

Jul 2020 – Jun 2021

**Core Courses:** Machine Learning, Natural Language Processing, Postgraduate Academic and Dissertation Writing – SE, Mathematical and Biological Foundation of Reinforcement Learning, Speech Processing

## PUBLICATIONS

[1] Cunxiang Wang\*, **Pai Liu\*** (equal contribution), Yue Zhang.

*Can Generative Pre-trained Language Models Serve as Knowledge Bases for Closed-book QA?*

The Joint Conference of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (**ACL-IJCNLP 2021**), long. [\[arxiv\]](#)

[2] **Pai Liu**

*QiaoNing at SemEval-2020 Task 4: Commonsense Validation and Explanation system based on ensemble of language model*

In Proceedings of The 14th International Workshop on Semantic Evaluation(**SemEval 2020**). [\[arxiv\]](#)

[3] Zhixiang Chen\*, Yikun Lei\*, **Pai Liu\*** (equal contribution), Guibing Guo

*NEUer at SemEval-2021 Task 4: Complete Summary Representation by Filling Answers into Question for Matching Reading Comprehension*

In Proceedings of The 15th International Workshop on Semantic Evaluation(**SemEval 2021**). [\[arxiv\]](#)

## RESEARCH EXPERIENCES

### Simple KBQA & Document-enhanced KBQA

Jul 2021 – Present

Research assistant | Westlake University | Supervisor: **Guibing Guo**

Hangzhou, China

- Proposed a domain-aware model to assist predicate matching in fact selection to pick the most relevant fact.
- Achieved state-of-the-art against all simple KBQA on SimpleQuestions and WebQuestionsSP dataset
- Adopted additional documents as an information source for background knowledge regarding the entities mentioned in the question
- Achieve state-of-the-art against all document-enhanced KBQA on WebQuestionsSP

### Emotion Causal Recognition

Apr 2021 – Present

Teach Assistant | Westlake University | Supervisor: **Yue Zhang & Libo Qin**

Hangzhou, China

- Proposed to predict the emotion of each sentence to assist the emotion causal recognition task, and to use the finer granularity of utterance-level context information interaction
- Proposed to use emotional information to shorten the distance of long-distance causal relationship pairs on RECCON dataset by constructing emotion based adjacent matrix and using GAT to exchange historical information
- Reduced the time complexity from  $O(mN)$  to  $O(N)$  compared with baseline by using a parallel prediction approach

### Relation Extraction

Feb 2021 – Present

Research Assistant | Westlake University | Supervisor: **Yue Zhang**

Hangzhou, China

- Applied BART+GDPNet to in the relation extraction task, which was the first attempt to the best of out knowledge
- Added coreference information annotated by AllenNLP to DocRED to transform partial inter-sentence relations into intra-sentence relations and improve F1 value by 0.4 points on ATLOP-Roberta-large

### Can Generative Pre-trained Language Models Serve as Knowledge Bases for Closed-book QA?

Jul 2020 – Feb 2021

Visiting Student | Westlake University | Supervisor: **Yue Zhang**

Hangzhou, China

- Proposed to apply SQuAD on closed-book QA, which contains corresponding passages, lower test-train overlap
- Explored the capability of generative language model to serve as KB for closed-book QA under a more strict and solid experimental setting.
- Proved that the reason for BART's poor performance on KBQA is its poor ability to memorize knowledge and extract internal knowledge of the model to answer questions

### **Covid-2019 Diagnosis Dialogue System**

*Mar 2020 – May 2020*

Remote Research Intern | University of California, San Diego | Supervisor: **Pengtao Xie**

Wuhan, China

- Conduct a background investigation on sequence-to-sequence dialogue system models
- Annotate data for the online diagnostic dialogue system

*(Forced to suspend cooperation because of Trump's ban)*

### **Commonsense Validation and Explanation**

*Mar 2020 – May 2020*

Independent Study | Northeastern University | No Supervisor

Wuhan, China

- Complete subtasks A and B of SemEval Task 4 with different language models and ensemble model, ranked 6<sup>th</sup>/45 and 11<sup>th</sup>/35, respectively

### **Text Summarization Incorporating reader comments & Rock Failure Prediction**

*Sep 2019 – Mar 2020*

Research Assistant | Northeastern University | Supervisor: **Guibing Guo**

Shenyang, China

- Proposed CommentPGNet that utilizes comments to create reader aware model, which helps the original article better summaries and increase accuracy by 5%
- Incorporated acoustic emission, stress and other information to predict the degree of rock damage using FM, LSTM and CNN

### **Language Parsing**

*Jun 2019 – Sep 2019*

Summer Intern | Northeastern University/NiuTrans | Supervisor: **Tong Xiao**

Shenyang, China

- Implement a named entity recognition system using dilated convolutional neural network + CRF architecture

## **TEACHING**

### **Teaching Assistant**

*Apr 2021- Jul 2021*

School of Engineering, Westlake University

*Natural Language Processing*

## **AWARDS AND HONORS**

- **2021 Outstanding Graduate** *(Awarded to top 3% of students)*
- **2021, 2019, 2018 The First Prize Scholarship of Northeastern University** *(Awarded to top 1% students)*
- **2019, 2018 National Scholarship** *(Awarded to top 0.5% students)*
- **2019, 2018 Academic excellence award of Northeastern University** *(Awarded to top 3% of students)*
- **2019 Excellence Award of Summer Internship Program, NiuTrans** *(Awarded to the best intern)*

## **INTEREST & LEADERSHIP**

### **Interests: fitness, badminton, popping**

- 2020 Champion of the 0-74kg class powerlifting of Westlake University
- 2019, 2018, 2017 Champion of badminton singles in software college at Northeastern University
- Vice President of Street Dance Club at Northeastern University (over 2,00 members)
- President of Street Dance Club at Westlake University (over 2,00 members)

## **SKILLS**

- **Programming Language:** *Python, Java, C, SQL, JavaScript, HTML*
- **Deep learning framework:** *Pytorch, Tensorflow*