

# PEI CHEN

<https://brickee.github.io/>

+1 979-739-4507 ◇ chenpei.net@gmail.com

## EDUCATION

---

### Ph.D. Candidate in Computer Science

2019.8 - 2023.12 (expected)

- Research Areas: Natural Language Processing;
- Research Interests: Large Language Model Pretraining and Prompting, Information Extraction;
- Overall GPA: 4.0/4.0 till now.

*Texas A&M University*

### MS in Finance

2016.9 – 2018.6

- Thesis: Does News Sentiment Predict the Stock Market? An Example on Chinese Growth Market;
- Received 2017 National Scholarship for Graduate Student;
- Overall GPA: 3.9/5.0, ranking 1/178.

*Southwestern University of Finance and Economics*

### B.Engr. in Simulation Engineering

2010.9 – 2014.6

- Thesis: Analyze and reconstruct the multi-resolution modeling technology of a simulation system written by millions of lines of C++ code;
- Overall GPA 88.61/100, ranking 1/45.

*National University of Defense Technology*

## EXPERIENCE

---

### Open Source AI Team, AWS AI

2023.5 – 2023.8

*Applied Scientist (Intern)*

*Santa Clara, CA*

Doing research on designing cutting-edge automatic prompting methods for large language models.

### Bedrock Team (Amazon Titan Model), AWS AI

2022.6 – 2023.1

*Applied Scientist (Intern)*

*Santa Clara, CA*

Pretrained large language models in AWS clusters; proposed a novel tabular language model that models tables as hypergraphs and learns better table representations for improving table understanding tasks.

### Tencent AI Lab

2021.6 – 2021.8

*NLP Researcher (Intern)*

*Seattle, WA*

Proposed a comprehensive benchmark for zero-shot knowledge base completion (KBC) tasks, covering state-of-the-art KBC methods and broad knowledge source data.

### Department of Computer Science & Engineering, Texas A&M University

2019.9 – 2023.5

*Research Assistant and Teaching Assistant*

*College Station, TX*

Research: Improved domain-specific named entity recognition task by modeling non-sequential entity mention relations using Graph Neural Networks; Improved fine-grained opinion mining task; Built fine-grained named location recognition benchmark, etc.

Teaching: CSCE 636 Deep Learning; CSCE 313 Computer Systems

### National Lab of Pattern Recognition, Chinese Academy of Sciences

2018.1 – 2019.8

*Research Engineer*

*Beijing, China*

Improved the event extraction and causality detection tasks from financial domain texts.

**Innovation Lab of Global Exchange, State Street**

*Data Analyst (Intern)*

2017.7- 2018.1

*Hangzhou, China*

Working on data cleaning, analysis, visualization and database construction for innovative financial applications.

## PUBLICATIONS

---

**Pei Chen**, Soumajyoti Sarkar, Leonard Lausen, Balasubramaniam Srinivasan, Sheng Zha, Ruihong Huang, and George Karypis. “[HYTREL: Hypergraph-enhanced Tabular Data Representation Learning.](#)”, 2023.

**Pei Chen**, Wenlin Yao, Hongming Zhang, Xiaoman Pan, Dian Yu, Dong Yu, and Jianshu Chen. “[ZeroKBC: A Comprehensive Benchmark for Zero-Shot Knowledge Base Completion.](#)” ICDM-2022, KG workshop.

**Pei Chen**, Haotian Xu, Cheng Zhang, and Ruihong Huang. “[Crossroads, Buildings and Neighborhoods: a Dataset for Fine-grained Location Recognition](#)”. NAACL-2022, long paper, acceptance rate: 21.96%.

**Pei Chen**, Haibo Ding, Jun Araki, and Ruihong Huang. “[Explicitly Capturing Relations between Entity Mentions via Graph Neural Networks for Domain-specific Named Entity Recognition.](#)” ACL-2021, short paper, acceptance rate: 21.2%.

**Pei Chen**, Kang Liu, Yubo Chen, Taifeng Wang, and Jun Zhao. “[Probing into the Root: A Dataset for Reason Extraction of Structural Events from Financial Documents.](#)” EACL-2021, short paper, acceptance rate: 24.7%.

**Pei Chen**, Hang Yang, Kang Liu, Ruihong Huang, Yubo Chen, Taifeng Wang, and Jun Zhao. “[Reconstructing Event Regions for Event Extraction via Graph Attention Networks.](#)” AACL-2020, long paper, acceptance rate: 28.3%.

## SKILLS

---

Competent: Python, PyTorch, PyTorch Lightning, Git

Familiar: C/C++, TensorFlow, Keras, SQL

Experienced: Parallel pretraining and training with multiple GPUs and machines.

## PROFESSIONAL SERVICE

---

2023: Program Committee/Reviewer for ACL, EMNLP

2022: Program Committee/Reviewer for EMNLP, ACL Rolling Review, NLPCC

2021: Program Committee/Reviewer for EMNLP, ACL Rolling Review, NLPCC