Yusen Zhang

1750 Clairmont Rd 29 yzh2749@emory.edu
Decatur, GA 30033 (404) 823-7578

EDUCATION

Emory University May 2021
Master of Science in Computer Science (GPA 3.85/4.0)
Beijing Institute of Technology June 2018
Bachelor of Science in Computer Science (ranked 16/166, GPA 3.6/4.0)
Relevant Coursework (score): CS557 Artificial Intelligence (102%), CS526 Algorithm (95%), CS534 Machine Learning

PROGRAMMING SKILLS

Proficient in ***Python, C, C++, Pytorch, Sklearn, etc.***

RESEARCH EXPERIENCE

Research on CONQUER for Nurses, Emory University, *Emory NLP Research Lab.* In progress
- Analyzed relation-aware self-attention model for text-to-SQL on Spider 1.0 Challenge, i.e., RATSQL etc.
- Currently, work on optimizing the Schema Encoder as well as Decoder by token level prefix tree and attention.

Research on Text-to-SQL, Yale University, *LILY Lab.* Jun. 2020 - Now
- Proposed the first cross-domain text-to-SQL question intention classification benchmark, TriageSQL. It requires models to distinguish four types of unanswerable questions from answerable questions, allowing the system to response differently according to the type of input. Submitted a short paper to NAACL 2021.
- Proposed a neural model for logic-consistent text generation from semantic parses, e.g. SQL-to-text. Also proposed an auto evaluation metrics to evaluate the logic-consistency of NL and logic form. Submitted a paper to NAACL 2021.

Research on Name Entity Recognition, Microsoft Research Asia, *Big Data Mining Group.* Jan. 2018 - Jun. 2018
- Proposed a novel CNN-based network for NER: gated relation network(GRN). The model could capture long-term context without using RNN layer and enjoyed lower time cost to train and test.
- Implemented proposed model GRN using Pytorch, then tested the model on two benchmark NER datasets, i.e. CoNLL03 and Ontonote5.0. GRN kept state-of-the-art on both datasets before BERT was proposed.

PROGRAMMING PROJECTS

Pattern System: A pattern-based intention recognition system, Microsoft Research Asia Oct. 2017 - Jan. 2018
- Implemented a domain-specific intent recognition platform that helps users build rule-based patterns to extract needed entities/intents. (lines of code: 27,000, hand-written by me: 17,000, certification available)
- Developed the system in C#.NET MVC/JavaScript based on the source code of Microsoft LUIS (Language Understanding Intelligent Service).
- Revised a human-computer interaction mechanism to enable users to mark utterances and store entities efficiently.

EdiLU: A DNN based entity extraction system, Microsoft Research Asia Jan. 2018
- Rewrote the original LSTM-Chain-CRF model (based on C++) code by CNTK (Microsoft Cognitive Toolkit, C# based), processed and trained eight benchmarks on NVIDIA Tesla P100 to evaluate the model.
- Help the engineering teams to smoothly encapsulate the rewritten model into Microsoft Office Software.

PUBLICATIONS

Chen, H.*, Lin, Z., Ding, G., Lou, J., Zhang, Y.*, & Karlsson, B. (2019, July). GRN: Gated relation network to enhance convolutional neural network for named entity recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, pp. 6236-6243).

Zhang Y.*, Dong X.*, Chang S., Tao Y., Rui Z. (2020, Oct). Did You Ask a Good Question? A Cross-Domain Question Intention Classification Benchmark for Text-to-SQL. In *Int-Ex Workshop at EMNLP 2020*.

**indicate students*

SELECTED COMPETITION AWARDS AND HONORS

- The 28th place(19 teams advanced to World Final), ACM-ICPC North America Champions Feb.2020
- The 3rd place, ACM-ICPC Southeast American Regional Contest Nov. 2019
- Award of Excellence, "Star of Tomorrow Internship Program" in Microsoft Research Asia Jun. 2018
- Science and Technology Progress Award, Beijing Institute of Technology (19/3000) Jan. 2018

ADDITIONAL INFORMATION

Services: Program Committee of InTex-SemPar @ EMNLP 2020
Languages: Fluent in written and conversational English and Chinese