YUWEI WU

Department of Computer Science and Engineering
School of Electronic Information and Electrical Engineering, Zhiyuan College
Shanghai Jiao Tong University, P.R. China
+86 13851805239 | will8821@sjtu.edu.cn
Homepage | GitHub

EDUCATION

Shanghai Jiao Tong University Shanghai, China
B.Sc. in Computer Science September 2016 – Present

- GPA: 3.5/4.0
- Selected award: KoGuan Encouragement Scholarship (top 1.5% of ~800 participants, awarded to those with the most outstanding academic performance out of all of the majors at SJTU)
- Selected to ACM Honored Class (an elite CS program at SJTU for the top 5% of students of Computer Science major)
- Visiting student scholar completing research on dialog systems at the Georgia Institute of Technology (July 2019 to December 2019)

PUBLICATIONS

1. "I Know What You Want: Semantic Learning for Text Comprehension" Zhuosheng Zhang, **Yuwei Wu**, Zuchao Li, Shexia He, Hai Zhao. Accepted by **PACLIC 33 (2019)** (**oral**) [Paper]

2. "Auto-Retoucher: A Framework for Background Replacement and Foreground Adjustment" Yunxuan Xiao, Yikai Li, **Yuwei Wu**, Lizhen Zhu. Accepted by **MVA 2019** (**oral**) [Paper]

3. "SG-Net: Syntax-Guided Machine Reading Comprehension" Zhuosheng Zhang*, **Yuwei Wu***, Junru Zhou, Sufeng Duan, Hai Zhao. Submitted to **AAAI 2020** [Paper]

4. "Semantics-aware BERT for Language Understanding" Zhuosheng Zhang*, **Yuwei Wu***, Zuchao Li, S Zhang, Xi Zhou and Xiang Zhou. Submitted to **AAAI 2020**

5. "Dual Co-Matching Network for Multi-choice Reading Comprehension" LiangShuai Zhang, Hai Zhao, **Yuwei Wu**, Zhuosheng Zhang, Xi Zhou, Xiang Zhou. Submitted to **AAAI 2020** [Paper]

*Equal authorship

RESEARCH EXPERIENCE

SJTU, BCMI Laboratory Shanghai, China
Research Assistant to Professor Hai Zhao June 2018 – June 2019
Semantic Role Labeling and Syntax for Text Comprehension

- Based on the observation that current language models generate semantically incomplete answers when answering questions, we proposed that semantics and syntax should be incorporated into model training as external knowledge.
- Our proposed model SemBERT has obtained new SOTA or substantially improved results on ten reading comprehension and language inference tasks. SG-Net has gained a significant improvement over the fine-tuned BERT baseline.
- The following papers have been submitted to AAAI 2020: "SG-Net: Syntax-Guided Machine Reading Comprehension" and "Semantics-aware BERT for Language Understanding."

Georgia Institute of Technology, Diyi Yang's Laboratory Atlanta, Georgia
Research Assistant to Professor Diyi Yang July 2019 – December 2019
Personalized Response Generation

- Due to a lack of high-quality personalized dialog datasets, we proposed our own dialog datasets, which consisted of 2 million conversations and 300,000 authors with extra author profile information and comment histories. We also proposed two dialog models and evaluation metrics to deal with personalized response generation.

COURSE PROJECTS

- MX-Compiler (Java) [Link] 2018
  - Compiled C-and-Java-like language to x86-64 Assembly in NASM assembly.

- Implemented graph coloring register allocation, semantic analysis, dead code elimination, and other optimizations such as function inlining and simple constant propagation.
- My compiler runs faster than gcc O1 on test set.

- SimpleDB-2019 (Java) [Link] 2019
  - Simple Database Management System in Java.
  - Implemented query optimization, locking, transactions, and concurrent queries.
- RISC-V CPU (Verilog) [Link] 2018
  - CPU with five-stage pipeline and two-way cache from scratch.
- Text Classification and Sentiment Analysis (Python) [Link] 2018
  - A binary classification task on News and Media Article
  - Won fourth place on the Kaggle leaderboard.
  - Course grade: 96/100 (top 5% of the 180 students who took the CS420 Machine Learning Course at SJTU)

TEACHING EXPERIENCE

CS 122: Computer Programming, Shanghai Jiao Tong University
Teaching Assistant September – December 2017

- Responsibilities included running recitation sections and designing and grading homework assignments

EXTRACURRICULAR ACTIVITIES AND SELECTED PROJECTS

Stanford **Qu**estion **A**nswering **D**ataset **(SQuAD) 2.0** Leaderboard on Machine Reading Comprehension July 2019

- The **best** in all single models and **2nd best** in all models in terms of EM and F1 scores
- The **first** team to surpass human benchmark on both EM and F1 scores with a single model in terms of submission time
- The **first** team to exceed 90% F1 score with ensemble models in terms of submission time
- [Leaderboard] [Paper] [Report]

ReA**d**ing **C**omprehension dataset collected from English **E**xaminations (**RACE)** Leaderboard on Machine Reading Comprehension March 2019

- The **best** among all submissions in terms of accuracy score
- [Leaderboard] [Paper] [Report]

Stanford **N**atural **L**anguage **I**nference **(SNLI)** Leaderboard on Language Inference April 2019

The **best** among all submissions in terms of accuracy score

- [Leaderboard] [Paper]

General **L**anguage **U**nderstanding **E**valuation **(GLUE)** Leaderboard on Language Understanding March 2019

The **3rd best** among all submissions in terms of averaged scores on eleven datasets

- [Leaderboard] [Paper]

SELECTED AWARDS AND HONORS

- Zhiyuan Honorary Scholarship (top 10% of all students of all majors from Zhiyuan college. Awarded to students with outstanding academic performance in their major). 2016, 2017, 2018
- Shanghai Jiao Tong Scholarship (top 20% of all students from all majors at Shanghai Jiao Tong University. Awarded to those with outstanding academic performance in their major). 2016, 2017, 2018
- Meritorious Winner of the Mathematical Contest in Modeling (top 10% of ~12000 international participants. Awarded to those with outstanding problem solving and writing skills). 2018

ADDITIONAL INFORMATION

Research Interests:

- Natural Language Processing, Machine Reading Comprehension, Question Answering
- Computational Social Science, Dialog Systems, Response Generation