

Justin Yang

DATA SCIENTIST · NLP RESEARCHER

✉ zake7749@gmail.com | 🏠 zake7749.github.io | 📞 zake7749 | 📱 kai-chou-yang

Summary

As a winner of many international data science competitions, I am experienced in data science, machine learning and widely used numerical libraries. I am also a NLP researcher, currently studying natural language understanding and sentence pair modeling. I've developed several NLP/IR related projects and published my researches on international conferences.

Honors

Cross-lingual Short-text Matching of Question Pairs, 2nd

CIKM'18 AnalytiCup

SINGLE COMPETITOR

Aug. 2018

- Proposed two densely-connected architectures for sentence pair modeling.
- Fused the features from different levels to create diversity intra-models.
- For more details, please refer to my slides, document and Github repository

Fake News Classification, 3rd

WSDM'19 Cup

TEAM LEADER

Feb. 2019

- Implemented various NLI networks and inject the world knowledge using BERT.
- Proposed a rumor-aware model based on the concept of decomposable attention.
- For more details, please refer to my slides, paper and Github repository

Gendered Pronoun Resolution Competition, 4th

Google AI

SINGLE COMPETITOR

Apr. 2019

- Conducted extensive analysis among hidden layers of BERT
- Proposed a multi-head siamese scorer for answer selection
- For more details, please refer to my solution overview. Paper for ACL is under review.

Experience

WMMKS Lab (Web Mining and Multilingual Knowledge System Lab)

Tainan, Taiwan

RESEARCH ASSISTANT

Jul. 2016 - Feb. 2017

- Developed a Chinese chatbot engine where people can define dialogue scenarios online.
- Implemented deep learning based, information retrieval based and template based replying APIs.
- Created an end-to-end workflow which included data crawling, text cleaning and sentence modeling.

Education

NCKU (National Cheng Kung University)

Tainan, Taiwan

M.S. IN COMPUTER SCIENCE AND ENGINEERING

Sep. 2017 - Exp. Aug. 2019

- GPA 4.30 / 4.30
- As a teaching assistant for Introduction to Data Science, Data Mining.
- As a lecturer for ML courses to AI Cup, Delta Electronics, Industrial Technology Research Institute, etc.

NCKU (National Cheng Kung University)

Tainan, Taiwan

B.S. IN COMPUTER SCIENCE AND ENGINEERING

Sep. 2013 - June. 2017

- GPA 3.92 / 4.30
- Honorable mention on the graduation exhibition
- Academic Excellence Awards 2016 and 2017
- Receive scholar grants from the Ministry of Science and Technology with a question answering system project

Open Source Project

Mianbot

Github

A FLEXIBLE CHOTBOT ENGINE BASED ON KEYWORD VECTOR MATCHING

- Got 600+ stars, 200+ forks on Github.
- Implemented the hierarchical fuzzy matching with word2vec.
- Support chit-chat and allow user to define customize conversation scenarios with JSON.
- Integrated with my other two open-source projects, PTT Crawler and PTT Chat Generator