Justin **Yang**

Data Scientist · NLP Researcher

✉ zake7749@gmail.com | ⌂ zake7749.github.io | ⌨ zake7749 | in kai-chou-yang

Sum**mary**

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.

Hon**ors**

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

CIKM'18 AnalytiCup

Single Competitor

Aug. 2018

- Proposed two densely-connected architectures for sentence pairs 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-awarded 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.

Exp**erience**

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.

Edu**cation**

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 - Exp. June. 2017

- GAP 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

Ope**n 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