

Da-Wei Lee

DATA & APPLIED SCIENTIST · AI/DATA/BACK-END

☎ (+86) 156-5225-0910 | ✉ dwlee@pku.edu.cn | 💻 daviddwlee84 | 🎓 Da-Wei Li



Summary

I'm a Data & Applied Scientist at Microsoft. Including internship in MSRA and STCA, I have been worked in Microsoft for more than 3 years. Enjoy being a **maker** - think of any creative idea and try hard to make it come true. A Vim addict, loved discovering new technologies, and problem-solving. Familiar with the following areas: AI Full-stack Development, NLP and Deep Learning, Recommender System, Big Data Processing and Analysis, and Back-end Development. Love open source, have 4 repositories with more than 30 stars and 1 repository with more than 130 stars.

Experiences

Microsoft STCA Bing Multimedia

Suzhou, Jiangsu, China

DATA & APPLIED SCIENTIST

Jul. 2021 - Present

- Provide better video recommendation qualities to increase user click-through rates and daily active users on different scenarios like Bing, MSN, and Edge.
- Shipped **Relevance Model** to control the relevance quality while increasing coverage, which has **2.2% defect rate reduction**, in several related video scenarios like MSN Article, Bing Super Caption, and Edge Underside.
- Shipped **CTR Model** on MSN Article with **Video-CI/UU +2.4%** and **Video-CI Ratio +2.33%** on 37 markets.
- To *increase development efficiency*, made an online **Related Video Debug Tool** and general utilities such as general data processing modules, fully automatic annotation pipeline, config visualization script.

Microsoft STCA Bing NLP Carina

Beijing, China

ALGORITHM INTERN

Jul. 2020 - Jun. 2021

- Built a **Numeric Information Extraction System** for financial reports. Designed annotation guideline, trained MRC-based Sequence Labeling, and built back-end API services with docker containers.
- Built models of the **AI Writer** for automatic article generation and rewriting by generation models and information retrieval methods.

Microsoft Research Asia Knowledge Computing

Beijing, China

RESEARCH INTERN

Dec. 2019 - May. 2020

- Academic research on how to generate slides from academic papers. Got published to AAAI 2021, and do a live sharing on Microsoft's Bilibili.

Publications

Towards Topic-Aware Slide Generation For Academic Papers With Unsupervised Mutual Learning

AAAI 2021 (CCF A)

DA-WEI LI, DANQING HUANG, TINGTING MA, CHIN-YEW LIN

May. 2021

- Generating slides from papers by extractive summarization techniques and unsupervised mutual learning.
- During mutual learning, two extractors (a Log-Linear Classifier and a Neural Sentence Selection Model) learn collaboratively by teaching each other and bootstrap their performances.

Open Relation Extraction with Non-Existent and Multi-Span Relationships

KR 2022 (CCF B)

HUIFAN YANG, DA-WEI LI, ZEKUN LI, DONGLIN YANG, BIN WU

Feb. 2022

- Designed Query-based Multi-head Open Relation Extractor (QuORE) to extract single/multi-span relations and detect non-existent relationships.
- QuORE is a multi-head framework construct with two sub-modules, the SSE (Single-Span Extraction) and the QASL (Query-based Sequence Labeling) which share loss during training and a selector to choose which result to use during inference.

Open Relation Extraction via Query-Based Span Prediction

KSEM 2022 (CCF C)

HUIFAN YANG, DA-WEI LI, ZEKUN LI, DONGLIN YANG, JINSHENG QI, BIN WU

May. 2022

- Designed Query-based Open Relation Extractor (QORE), a multilingual Transformers-based language model to derive a representation of the interaction between arguments and context.
- Experiments including Multilingual ORE, ablation study on non-query/query-based and non-LM/LM models, zero-shot domain transferability, and few-shot learning ability.

Education

Peking University

Beijing, China

MASTER IN SOFTWARE ENGINEERING

Sep. 2018 - Jul. 2021

- Focus on NLP and related application in Knowledge Graph.
- Core member of PKU Open Source Association in ML/NLP department.
- Do research of Knowledge Extraction, Knowledge Graph, and Participating cases of Anti-healthcare fraud and Medical record analysis in Peking University National Engineering Research Center of Software Engineering.

Taiwan Tech (National Taiwan University of Science and Technology)

Taipei, Taiwan

BACHELOR IN ELECTRONIC AND COMPUTER ENGINEERING (MINOR FINANCE)

Sep. 2013 - Jun. 2017

- Focus on Embedding System Design but also App, Web, Game development.
- Got 4 individual piecework experiences and 5 engineering competitions based on different topics.