

XINRAN ZHAO

Email: xzhaoar@connect.ust.hk | GitHub: colinzhaoust | Website: colinzhaoust.github.io

EDUCATION

The Hong Kong University of Science and Technology

2016 - 2020, Hong Kong

- GPA = 3.958 (Top 1%, Academic Achievement Medal).
- Bachelor of Engineering, majoring in Computer Science.
- First Class Honors, Continuous Undergraduate Scholarship, Dean's list for the active semesters.
- Relevant Coursework:
 - Explore & Visualize Data; OOP & Data Structures (Honor); Discrete Math Tools; Design & Analy. of Algori; Statistical Learning Models for Text and Graph Data (Graduate Level); Big Data Mining & Management; Deep Learning for Computer Vision; Matrix Algebra

Cornell University

2018, NY, US

- International Exchange Program in Computer Science.
- Relevant Coursework: Machine Learning; Natural Language Processing; Operating Systems

RESEARCH EXPERIENCE

WinoWhy: A Deep Diagnose of Essential Commonsense Knowledge for Answering Winograd Schema Challenge

In **ACL 2020**, advised by Hongming Zhang and Prof. Yangqiu Song

HKUST, HK

- Proposes a new task, WinoWhy, which requires models to determine the plausibility of collected human reasoning from similar adversarial examples from humans and generative models. The task based on commonsense pronoun coreference choice questions further examines the model performance of understanding the commonsense reasoning instead of solely making correct prediction. Experiments on categorized human reasoning point out the limitation and possible future directions of current models.
- Personal Contribution: Proposed the idea in discussion; collected and rectified the dataset; implemented the models and baselines; wrote and reviewed the draft with collaborators.
- Paper was accepted by ACL 2020 and available on this [link](#).

PCR4ALL: A Comprehensive Evaluation Benchmark for Pronoun Coreference Resolution

In submission, advised by Hongming Zhang and Prof. Yangqiu Song

HKUST, HK

- Proposes a novel benchmark, PCR4ALL, which bridges the gap between document-level (e.g., CoNLL-2012) and sentence-level (e.g., Winograd Schema Challenge) pronoun coreference questions to evaluate the systems from different angles (i.e., knowledge sources, domains, frequency, bias) to check their reliability in real applications. Experiments demonstrate a notable gap between existing PCR models and a reliable PCR system since they are often optimized on a single target type and lack generalizability.
- Personal Contribution: Proposed the idea in discussion; implemented the code for dataset alignment and data labeling; implemented the models and baselines; wrote and reviewed the draft with collaborators.
- Paper to be submitted to TACL in 2021.

Exploring the Role of Debate Topic for Argument Persuasion

In submission, advised by Esin Durmus, Hongming Zhang, and Prof. Claire Cardie

Cornell, NY

- Proposes a bi-encoder model consisting of separate encoders for (1) the background and linguistic features of debaters ; (2) the language and the relations of the topics, to involve the information of debate topics to study more contextualized argument persuasiveness than the previous works that commonly focus on the debaters' characters and arguments alone.
- Personal Contribution: Proposed the idea in discussion; implemented the model and the baselines; performed qualitative and quantitative analysis; wrote and reviewed the draft with collaborators.
- Paper to be submitted to ACL 2021.

User Attitude Classification with Multitask Learning

In submission, advised by Esin Durmus and Prof. Claire Cardie

Cornell, NY

- Proposes a multitask learning (MTL) framework to help predict the online forum users' stance on controversial political issues with their background information and previous arguments through leveraging the possible correlations among the issues, which shows robust improvement than single-task baselines. Similar tasks predicted by a novel embedding-based method show great correlation with human intuition and sociological findings, with which the auxiliary task selection in MTL is automated and improved.
- Personal Contribution: Proposed the idea in discussion; implemented the model and the baselines; performed qualitative and quantitative analysis; wrote and reviewed the draft with collaborators.
- Paper to be submitted to ACL 2021.

A Brief Survey and Comparative Study of Recent Development of Pronoun Coreference Resolution

In ArXiv, advised by Hongming Zhang and Prof. Yangqiu Song

HKUST, HK

- Surveys in detail on the Pronoun Coreference Resolution (PCR) datasets (ordinary ones, hard ones requiring external knowledge, and others for special purposes) and models (from traditional ones to neural-based end-to-end models) commonly used for the task. Exhaustive experiments and comparisons show that some performance gain relies on the dataset similarity and knowledge graph based improved models relieve the problem of predicting infrequent cases.
- Personal Contribution: Implemented or replicated the models and the baselines; performed qualitative and quantitative analysis; wrote and reviewed the draft with collaborators.
- Paper arXived with code available on this [link](#) (code).

Learning Contextual Causality from Time-consecutive Images

In submission, advised by Hongming Zhang and Prof. Yangqiu Song

HKUST, HK

- Designs a novel task studying contextual causal knowledge from videos and proposed a corresponding dataset containing human annotated events and causality inferred from video frame sequences. Then we propose a Vision-Contextual Causal (VCC) model that can utilize the images as context to better acquire causal knowledge from temporal consecutive videos.
- Paper was submitted to AAAI 2021 (under Phase II reviewing).

Seek to Embed ASER: A Large-scale Eventuality Knowledge Graph

In WWW 2020, advised by Xin Liu and Prof. Yangqiu Song

HKUST, HK

- Designs a novel embedding model combining text embedding and graph embedding algorithms to learn the node representation for ASER, which has text as nodes and eventuality relations as edges. Provides useful signals for link prediction, unknown event resolution, and representation captioning on the graph, with a model utilizing BERT embeddings and LSTM.
- Applied ASER to solve the commonsense reasoning problems (Section 6.2) and beat the state-of-the-art at that time.
- Paper was accepted by WWW 2020 and available on this [link](#) (as an acknowledged contributor).

An Online Learning Platform with NLP Supported Teaching Assistant

Final Year Project, advised by Prof. Dit-Yan Yeung

HKUST, HK

- Seeks to build an online learning platform with NLP supported widgets, including a question answering module that retrieves video chunks from lecture videos to answer students' queries and a quiz generator that automatically generates basic multiple choice quizzes from auto-rectified and masked lecture subtitles.

Mining Course Structure for Course Recommendation

Undergraduate Research Opportunity, advised by Prof. Raymond Wong

HKUST, HK

- Builds a course recommendation and academic planner model based on course dependency graph (generated from school major requirements) and past students' performance records. Serves as part of the unreleased official HKUST undergraduate course planner, which provides requirement fulfillment checker and study path recommendation service for university students.

A Comparative Study on the Sentimental Characteristics of Chinese and Western Tourists (in Chinese)

Under review, advised by Yue Xi and Prof. Yushi Jiang

Chengdu, China

- Builds an NLP supported model to detect the comparative behavior differences on sensitiveness and attentiveness of Airbnb tourists from various cultures. Paper was presented in the Annual Conference of JMS China Marketing Science 2019.

WORK EXPERIENCE

Faithful Evaluation Metric for Text Simplification

Research Assistant, supervised by Prof. Dit-Yan Yeung

Sep 2020 - present, HK

- Our primary goal is to build AI supported agents that could help online language learners. We seek to design a faithful evaluation metric for text simplification by involving robust semantic similarity checker and generalizable simplicity detector.

Narrow Band Communication

Engineer Intern, Sichuan Jinhutong Tech. Stock Company

June 2017 - Sep 2017, Chengdu

- Assisted in the development of a smart office system for local China Mobile Headquarters based on Narrow Band Communication, writing C++ for Integrated Circuit components.

SKILLS

Programming: Python, C++, Java, JavaScript, Matlab, Pytorch, Keras/Tensorflow, and etc.

Languages: English (Proficient, with TOEFL=112), Mandarin (Native).

Debate: HKUST Madarin Debate Team: Championship for Bayarea Debate Invitational; Top 8 for the Sixth and Seventh International Mandarin Debate Invitational in Singapore.

Interests: Archery, Gym, Engraving, Basketball, Traditional Poem Writing