XUEQING WU

(217)-305-0509 ◇ xueqing.wu@cs.ucla.edu ◇ shirley-wu.github.io ◇ google scholar

EDUCATION

University of California, Los Angeles 09/2023 - Present
PhD in Computer Science (in progress). Advisor: Kai-Wei Chang, Nanyun Peng.

University of Illinois Urbana-Champaign 08/2021 - 05/2023
MS in Computer Science. Advisor: Heng Ji.

University of Science and Technology of China 09/2016 - 06/2020
BS in Electronic Engineering and Information Science. GPA: 4.03/4.30 (rank: 1/363).

INTERNSHIP & RESEARCH

University of California, Los Angeles 09/2023 - Present
Research Assistant. Mentor: Kai-Wei Chang, Nanyun Peng

· Research automatic self-debugging for visual programming. (Work in progress)

Bytedance AI Lab 05/2023 - 09/2023
Research Intern. Mentor: Haoran Huang

· Proposed a new dataset DACO for *data analysis*, containing 440 diverse databases, ∼2k data with automatic annotations, and 100 data with human annotations. Proposed to generate high-quality answers automatically via multi-turn code generation. Designed the DACO-RL algorithm to train the data analysis model with dense reward and boosted answer helpfulness in 57.72% cases. [9]

University of Illinois Urbana-Champaign 08/2021 - 05/2023
Research Assistant. Mentor: Heng Ji

· Built a high-quality dataset for *open-vocabulary state tracking* via large-scale human annotation. Designed a cluster-based metric to prevent over-rewarding repetitive predictions. Proposed two techniques, *entity memory* and *entity-conditioned prediction*, to improve the performance. [2]

· Proposed and benchmarked a novel task of *cross-document misinformation detection*, including document-level and fine-grained *event-level* detection. Employed cross-document event coreference resolution to construct cross-document knowledge graph. Designed a graph neural network based detector that significantly outperforms existing methods by up to 7 F1. [3]

IBM Research 05/2022 - 08/2022
Research Intern. Mentor: Alfio Gliozzo

· Proposed a retrieval-augmented model for three *table augmentation* tasks including cell filling, row population and column population. Trained the model via self-supervision without relying on ground truth for retrieval. Achieved state-of-the-art on two datasets with absolute MRR gains of up to 30% compared to non-retrieval baselines. [1]

Bytedance AI Lab 07/2020 - 07/2021
Research Intern. Mentor: Hang Li

· Proposed and benchmarked *text-to-table*, a novel IE setting that extracts table-format information and requires no pre-defined schema. Adopted a seq2seq approach and proposed two techniques to improve table generation, which significantly outperforms named entity extraction and relation extraction. [4]

Microsoft Research Asia 10/2019 - 06/2020
Research Intern. Mentor: Tao Qin

- Proposed a new sequence learning framework which boosts a given main task using *temporally correlated tasks* as auxiliary training tasks. Designed a novel algorithm using RL to jointly train the task scheduler, and improved the baselines on four simultaneous translation tasks and a stock forecasting task. [5]
- Employed BERT as the contextual feature extractor for context-aware machine translation, and analyzed three feature aggregation methods. Achieved state-of-the-art on three datasets. [6]
- Developed a simple but effective data augmentation method for machine translation based on sentence concatenation and improved the performance on nine datasets. [7]

University of Science and Technology of China 03/2019 - 06/2019
Undergraduate Researcher. Mentor: Jun Du

- Proposed two techniques for oriented object detection in aerial images: a novel tilt angle representation to avoid ambiguity, and an IoU calculation method to increase positive candidates of long objects in RPN stage. Won the first prize (out of 32 participants) in ODAI Challenge-2019 held by CVPR. [8]

PUBLICATIONS

[1] Michael R. Glass, **Xueqing Wu**, Ankita Naik, Gaetano Rossiello, Alfio Gliozzo, *Retrieval-Based Transformer for Table Augmentation*, ACL findings. 2023. Link

[2] **Xueqing Wu***, Sha Li*, Heng Ji, *OpenPI-C: A Better Benchmark and Stronger Baseline for Open-Vocabulary State Tracking*, ACL findings (short paper). 2023. Link

[3] **Xueqing Wu**, Kung-Hsiang Huang, Yi Fung, Heng Ji, *Cross-document Misinformation Detection based on Event Graph Reasoning*, NAACL. 2022. Link

[4] **Xueqing Wu**, Jiacheng Zhang, Hang Li, *Text-to-Table: A New Way of Information Extraction*, ACL. 2022. Link

[5] **Xueqing Wu**, Lewen Wang, Yingce Xia, Weiqing Liu, Lijun Wu, Shufang Xie, Tao Qin, Tie-Yan Liu, *Temporally Correlated Task Scheduling for Sequence Learning*, ICML. 2021. Link

[6] **Xueqing Wu**, Yingce Xia, Jinhua Zhu, Lijun Wu, Shufang Xie, Tao Qin, *A Study of BERT for Context-Aware Neural Machine Translation*, ACML journal track. 2021. Link

[7] **Xueqing Wu**, Yingce Xia, Jinhua Zhu, Lijun Wu, Shufang Xie, Yang Fan, Tao Qin, *mixSeq: A Simple Data Augmentation Methodfor Neural Machine Translation*, IWSLT Workshop. 2021. Link

[8] Yixing Zhu, Jun Du, **Xueqing Wu**, *Adaptive Period Embedding for Representing Oriented Objects in Aerial Images*, IEEE Transactions on Geoscience and Remote Sensing. 2020. Link

PREPRINTS

[9] **Xueqing Wu**, Rui Zheng, Te-Lin Wu, Hanyu Zhou, Tang Mohan, Kai-Wei Chang, Nanyun Peng, Haoran Huang, *DACO: Towards Application-Driven and Comprehensive Data Analysis via Code Generation*, Arxiv 2403. Link, Website, Code

AWARDS & HONORS

Graduate Dean's Scholar Award, UCLA 09/2023
Siebel Scholar (awarded annually to over 90 top students from the world), UIUC 09/2022
First Prize and **Second Prize**, Challenge-2019 on Object Detection in Aerial Images 04/2019
Honorary Rank (top 5%), USTC 11/2019
Guo Moruo Scholarship (highest honor at USTC, top 1%), USTC 10/2019
Tang Lixin Scholarship, USTC 10/2018
China National Scholarship (top 0.2% of national candidates), China 10/2018