

Shih-Chieh Dai

+1(512)705-9690 | sjdai@utexas.edu | sjdai.github.io

RESEARCH INTERESTS

Natural Language Processing, Explainable AI, Human-centered NLP

EDUCATION

- The University of Texas at Austin**, M.S. in Information Science August 2021 – May 2023
Concentration: Applied Machine Learning Austin, TX
- National Chengchi University**, B.S. in Management Information Systems September 2016 – June 2020
Concentration: Natural Language Processing, Application Development Taipei, Taiwan

RESEARCH EXPERIENCE

Research Assistant. Penn State University & Academia Sinica June 2022 – Present
Advisor: Dr. Aiping Xiong & Dr. Lun-Wei Ku Remote

- Researched on **Chain-of-Thoughts** to thematic coding for open-ended responses.
- Conducted human-subject experiments with **1500+** participants recruited from **Amazon Mechanical Turk** and **Prolific**.
- Proposed a framework (**QA-model**, **entailment model** and **transformer model**) for generating the explanation of the fact-checked results with **70%** correctness. (Publication [C1])

Independent Study. Human-Centered Computing Lab, UT-Austin September 2021 – May 2022
Advisor: Dr. Matt Lease Austin, TX

- Researched on Human-In-the-Loop classification using **Active Learning** and rationales.
- Implemented a **BERT Transformer** to extract the rationales on the news data.
- Trained machine learning model using HPC scheduler, **Slurm**.

Research Assistant. Academia Sinica July 2017 – August 2021
Advisor: Dr. Lun-Wei Ku Taipei, Taiwan

- *Fake News Influence Mitigation*
 - * Created an interactive platform for non-technical users to visualize the stance of the news articles using **D3.js** and set up **Elasticsearch** as the search engine for **610,000+** news articles. (Publication [D1])
 - * Researched on **Reinforcement Learning** with delivering **6%** more verified news versus baseline. (Publication [C2])
 - * Developed a news article website using **Flask**, **JavaScript**, **MongoDB** to integrate a fake news intervention API and deploy it using **Docker**, **Nginx** with **172** visits. (Publication [D2])
- *NLP Applications*
 - * Applied a BERT-based model and built a web interface for English teachers to examine the quality of the test paper with **86.3%** accuracy. (Publication [D3])
 - * Built ETL pipelines for automated processing of Text and Speech data.

CONFERENCE PUBLICATIONS

(* indicates contributed equally)

- [C1] **Shih-Chieh Dai***, Yi-Li Hsu*, Aiping Xiong, Lun-Wei Ku, “Ask to Know More: Generating Counterfactual Explanations for Fake Claims”, In: *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*. 2022 (**Oral Presentation**)
- [C2] Kuan-Chieh Lo, **Shih-Chieh Dai**, Aiping Xiong, Jing Jiang, Lun-Wei Ku, “VICTOR: An Implicit Approach to Mitigate Misinformation via Continuous Verification Reading”, In: *ACM The Web Conference (WWW)*. 2022
- [C3] Yi-Pei Chen, Ting-Lun Hsu, Wen-Kai Chung, **Shih-Chieh Dai**, Lun-Wei Ku, “Upstream, Downstream or Competitor? Detecting Company Relations for Commercial Activities”, In: *International Conference on Human-Computer Interaction (HCII)*. 2019 (**Oral Presentation**)

SYSTEM DEMONSTRATIONS

- [D1] **Shih-Chieh Dai***, Kuan-Chieh Lo*, Aiping Xiong, Jing Jiang, Lun-Wei Ku, “*Escape from Echo Chamber*”, In: *ACM The Web Conference (WWW)*. 2021
- [D2] **Shih-Chieh Dai***, Kuan-Chieh Lo*, Aiping Xiong, Jing Jiang, Lun-Wei Ku, “*All the Wiser: Fake News Intervention Using User Reading Preferences*”, In: *ACM International Conference on Web Search and Data Mining (WSDM)*. 2021
- [D3] **Shih-Chieh Dai***, Ting-Lun Hsu*, Lun-Wei Ku, “*Exam Keeper: Detecting Questions with Easy-to-Find Answers*”, In: *ACM The Web Conference (WWW)*. 2019

WORK EXPERIENCE

- Student Technician | IT Ops Team** September 2021 – May 2023
University of Texas at Austin *Austin, TX*
- Utilized **Azure Lab Service** to set up virtual machines for professors to create lab assignments.
 - Developed a **Bash** script to automate the process of leaving and rejoining UT’s active directory domain.
 - Designed and implemented a CI/CD pipeline using **GitHub Actions** to deploy faculty member websites.
- Lab Infrastructure Manager/System Administrator** January 2019 – August 2021
Natural Language Processing and Sentiment Analysis Lab, Academia Sinica *Taipei, Taiwan*
- Manage 11 GPU servers with 22 GPUs in the lab and maintain web server, file server (NAS) and database
 - Wrote documentation and trained new lab members for their onboard process for using Lab’s computing resources.

TEACHING EXPERIENCE

- Teaching Assistant — Course : Programming 101** February 2019 – June 2019
Department of Computer Science, National Chengchi University *Taipei, Taiwan*
- Taught Python programming and guided the students to complete their capstone project

PROJECTS

- Grocery Expiration Reminder | React, Go, Postgres** March 2023 – Present
- Designed and built the responsive front-end interface using **React**
 - Implement the APIs using **Go** for accessing **Postgres** database.
- Patients Readmission Risk Predictions | Python, Scikit-learn** March 2022 – May 2022
- Implemented **Random Forest**, **Logistic Regression**, and **SVM** using scikit-learn for patients readmission prediction on Nationwide Readmission Database with **290,000+** records.
 - Extracted essential features using the Chi-Square test to improve **12%** accuracy scores.
- MindReader | Python, Scikit-learn, SQL server, MongoDB, Flask** January 2019 – December 2019
- Constructed a merchandise recommendation system by **Scikit-learn**.
 - Wrapped the recommendation system as RESTful API with **Flask**.

AWARDS AND GRANTS

- [a] **Travel Grant | \$2,000** 2022
- Attending KDD, Washington, D.C.
- [b] **Graduation Project Competition | Second Place** 2019
- MindReader-A Micro-moment Recommendation Mechanism API—(Advisor: Prof. Yi-Ling Lin)
- [c] **Travel Grant | \$3,000** 2019
- Attending The Web Conference (WWW), San Francisco, CA
- [d] **Travel Grant | \$3,000** 2019
- Attending International Conference On Human-Computer Interaction (HCII), Orlando, FL

SKILLS

Programming Languages: Java, Python, JavaScript, Go
Web Development: Flask, React, HTML/CSS, Nginx
Machine Learning: PyTorch, Keras, Scikit-learn, Pandas, Numpy, NLTK, GenSim, Huggingface
Database: MongoDB, Postgres, MySQL
Others: Azure, AWS, Git, Docker, Elastic Search, Vim, Linux, CI/CD, L^AT_EX