Shih-Chieh Dai

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

Research Interests

Natural Language Processing, Explainable AI, Human-centered NLP

EDUCATION

| The University of Texas at Austin, M.S. in Information Science Concentration: Applied Machine Learning | August 2021 – May 2023 Austin, TX | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|--|
| National Chengchi University, B.S. in Management Information Systems Concentration: Natural Language Processing, Application Development | September 2016 – June 2020 Taipei, Taiwan | |
| Research Experience | | |
| Research Assistant . Penn State University & Academia Sinica <i>Advisor:</i> Dr. Aiping Xiong & Dr. Lun-Wei Ku | June 2022 – Present Remote | |
| • Researched on Chain-of-Thought to thematic coding for open-ended responses | 3. | |
| • 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 <i>Advisor:</i> Dr. Matt Lease | September 2021 – May 2022 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 <i>Advisor:</i> Dr. Lun-Wei Ku | July 2017 – August 2021 Taipei, Taiwan | |
| • Fake News Influence Mitigation | | |
| * Created an interactive platform for non-technical users to visualize the stan D3.js and set up Elasticsearch as the search engine for 610,000+ news a * Researched on Reinforcement Learning with delivering 6% more verified (Publication [C2]) * Developed a news article website to integrate a RL verified news recommended | ce of the news articles using rticles. (Publication [D1]) I news versus baseline. der that outperforms the | |
| baseline system by 9% in terms of helpfulness in debunking fake news. (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
- [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

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 University of Texas at Austin Utilized Azure Lab Service to set up virtual machines for professors to creat Developed a Bash script to automate the process of leaving and rejoining UT's Designed and implemented a CI/CD pipeline using GitHub Actions to deplot Lab Infrastructure Manager/System Administrator Natural Language Processing and Sentiment Analysis Lab, Academia Sinica Manage 11 GPU servers with 22 GPUs in the lab and maintain web server, file | September 2021 – May 2023 Austin, TX e lab assignments. s active directory domain. y faculty member websites. January 2019 – August 2021 Taipei, Taiwan s server (NAS) and database |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| • Wrote documentation and trained new lab members for their onboard process for resources. | for using Lab's computing |
| Teaching Experience | |
| Teaching Assistant — Course : Intro to Programming Department of Computer Science, National Chengchi University Taught Python programming and guided the students to complete their capsto | February 2019 – June 2019 <i>Taipei, Taiwan</i> ne project |
| Projects | |
| Grocery Expiration Reminder React, Go, Postgres Designed and built the responsive front-end interface using React Implement the APIs using Go for accessing Postgres database. | March 2023 – Present |
| Patients Readmission Risk Predictions Python, Scikit-learn Implemented Random Forest, Logistic Regression, and SVM using scikit-prediction on Nationwide Readmission Database with 290,000+ records. Extracted essential features using the Chi-Square test to improve 12% accurace | March 2022 – May 2022 learn for patients readmission y scores. |
| MindReader Python, Scikit-learn, SQL server, MongoDB, Flask Ja • Constructed a merchandise recommendation system by Scikit-learn. • Wrapped the recommendation system as RESTful API with Flask. | nuary 2019 – December 2019 |
| Awards and Grants | |
| [a] Travel Grant \$2,000Attending KDD, Washington, D.C. | 2022 |
| [b] Graduation Project Competition Second Place MindReader-A Micro-moment Recommendation Mechanism API—(Advisor: P | 2019 rof. Yi-Ling Lin) |
| [c] Travel Grant \$3,000 Attending The Web Conference (WWW), San Francisco, CA | 2019 |
| [d] Travel Grant \$3,000 Attending International Conference On Human-Computer Interaction (HCII), 0 | 2019 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, IATEX