Yuelyu Ji

Email: yueluji@gmail.com

LinkedIn: https://www.linkedin.com/in/yuelyu-ji-265495154/

EDUCATION

University of Pittsburgh — Pittsburgh, USA
PhD in Information Science — *Aug* 2023 - Present
University of Pittsburgh — Pittsburgh, USA
Master in Information Science GPA: 3.68 — Jan 2021 - Dec 2022
Nanjing Agricultural University (NJAU) — Nanjing, China
Bachelor in Information System GPA: 3.50 — Sep 2016 - Jun 2020

RESEARCH EXPERIENCE

NetEase Cloud Music | job title Machine learning intern March 2023 - July 2023

- Assisted in developing Continuous Parameterization for Controlled Text Generation (CPCTG), a new approach for controlled text generation, using a Large Language Model (LLM) for style learning and **Proximal Policy Optimization (PPO)** for balancing factual and emotional content.
    - Use the reinforcement learning method to the Rouge-L score from 43% to 60% at the text generation task.
    - Authored a paper on the project, contributing to the field of multi-modal generation by improving text alignment with the input and enhancing control over text quality.

Graduate Student Researcher, Information Retrieval Integration and Synthesis (iRiS) Lab, University of Pittsburgh Advised by Dr. Daqing He Aug 2021 - present

- Automatic Classification of ADRD Caregivers' Online Information Wants: A Machine Learning Approach.
    - Alzheimer's-related posts and comments from Reddit from 2010 to 2020 were collected and categorized according to the Health Information Want of Alzheimer's caregivers.
    - At the data processing and model level, we used the **GPT-2 based few-shot learning model** to classify all the sentences in 200 labeled posts, and then put the sentences containing question marks or question keywords as questions, and the remaining identified sentences as background into **SOTA's QA model** for pre-training, and obtained a 40% The result of the QA model is then used as a question. The results of the QA model are then fed into SOTA's few-shot text classification task as a summary of the corresponding posts, and the F1 score up to **69%**.

University of Pittsburgh Advised by Dr. Ye Ye May 2022 - present

- Transfer learning based Covid-19 Readmission Risk Prediction
    - Using the Covid-19 readmission patients Electronic Health Records in 16 different hospitals and using the NER model to extract the medical entities and match medical entities with unique identifier CUIs. Get most 1000 frequent CUIs and generate dataset.
    - Using Domain-Adversarial Training of Neural Networks (DANN) to predict different readmission rate in different hospitals and change the DANN as muti-source oriented task the AUROC result of the Multi-DANN up to 85% .

PUBLICATIONS

Journal publications

- **Yuelyu Ji**, Disheng Liu, Runxue Bao, Qi Li, Ye Ye, COVID19 Readmission Risk Prediction using Electronic Health Records, *Journal of the American Medical Informatics Association* (ICHI) accepted!