Yiyang Feng

yiyang.feng.2375@gmail.com | https://wind-2375-like.github.io

EDUCATION

École Polytechnique Fédérale de Lausanne (EPFL) Sep 2022-Present

Master's Student in Computer Science; Overall GPA: 5.74/6.00

Xi'an Jiaotong University (XJTU) Aug 2018-Jul 2022

Bachelor in Automation Science & Technology; Overall GPA: 4.00/4.30; **Major GPA: 4.11/4.30** (top 5% among 197 students)

SELECTED RESEARCH EXPERIENCE

Bachelor's Thesis: Cross Domain Chinese Speech-to-SQL System Design Dec 2021-Jun 2022

Research Assistant, Intelligent Human-Computer Interaction Lab

Advisor: Zhongmin Cai, Professor, Department of Automation Science and Technology, Xi'an Jiaotong University

- Proposed several optimization methods for applying current Text-to-SQL systems on Chinese datasets: translation-model-based schema-linking and meta learning for domain generalization; improved the validation accuracy at most by 6.6%
- Applied the optimization methods to TypeSQL, SyntaxSQLNet, and IRNet on CSpider and compared their performance
- Combined Chinese speech-to-text model with text-to-SQL systems to build a Chinese speech-to-SQL platform

Explainable Graph Neural Networks for NLP May 2021-Oct 2021

Research Assistant, PSU NLP Lab

Advisor: Rui Zhang, Assistant Professor, Computer Science and Engineering Department, Pennsylvania State University

- Explored explainable Graph Neural Network (GNN) methods on two NLP tasks: DialogueGCN and TextGCN
- Applied GraphMask and GNNExplainer to interpret edge importance on predictions from selected trained GNN models
- Masked 50%+ of messages, most from the bottom layer, with an accuracy drop of less than 4% for both tasks
- Investigated the problem from the experimental result and the limitation of reparameterization tricks from GraphMask

Pseudocode Programming Based on Dialogue System Jun 2020-May 2021

Research Assistant, Intelligent Human-Computer Interaction Lab

Advisor: Zhongmin Cai, Professor, Department of Automation Science and Technology, Xi'an Jiaotong University

- Devised a pseudocode rule base to define the syntax of atomic operations in pseudocodes
- Utilized a traditional lexical and syntax analyzer in pseudocode translation based on the rule base with Lex and Yacc
- Designed a dialogue system that throws exceptions in natural languages to help users translate pseudocode into C code

SELECTED PROJECTS

Fine-tuning and Prompt-learning on Commonsense Causal Reasoning (CCR) Nov 2022-Dec 2022

ML4Science Project for CS-433 Machine Learning, EPFL

- Derived a new cause/effect generation task from the original real/fake causal classification task in the COPA dataset
- Conducted experiments on both tasks using fine-tuning (BART, RoBERTa, ALBERT) and prompt learning models (GPT-3).
- Compare the performance of two sets of models and analyze the results using BLEU, METEOR, ROUGE-L and CIDEr metrics

Gender Lens - A Cinematic Analysis Oct 2022-Dec 2022

*Team (**All**datapoint**a**ccurate) Project for CS-401 Applied Data Analysis, EPFL*

- Proposed to analyze gender stereotypes in the CMU Movie Dataset, and developed the web interface and drafted the data story
- Processed the movie summary text to extract neighboring verbs and adjectives of the names of male and female characters
- Identified gender stereotypes from difference between two sets of word distributions, using bar plot, χ^2 test, and KL divergence
- Compared the distribution divergence across different genres and time intervals from 1913 to 2013 to observe its evolution

SKILLS

- **Programming:** Proficient with Python; Knowledge in C, C++, MATLAB, JavaScript, CSS, and HTML
- **Machine Learning and NLP tools:** Pytorch, Huggingface Transformers, scikit-learn, spaCy, NLTK, AllenNLP, Numpy