Xiang Liu

Email: xliu29@gmu.edu

Mobile: +852-97479004

Education

- **The University of Hong Kong** — HK, CHINA
  Master of Science(MSc) - Computer Science — *Sep 2022 - Dec 2023*
- **George Mason University** — VA, USA
  Bachelor of Science - Computer Science; GPA: 3.71/4.0 — *Aug 2018 - Aug 2022*
  Honors/Awards: *Dean's List (2018-2020)*
  Courses: *Natural Language Processing, Visual Computing, Data Structure, Database Concepts*
- **Henan University** — Henan, CHINA
  Bachelor of Engineering - Networking Engineering; GPA: 88/100; Rank: 4/292 — *Sep 2017 - Jun 2022*

Experience

- **Baidu Research Cognitive Computing Lab** — Beijing,CHINA
  Intern (Full-time) — *Dec 2021 - June 2022*
  - Worked on dependency parsing, using Open Information Annotation (OIA) method to parse sentences.The OIA method will convert sentence to a DAG.
  - Improved the performance of the OIA method for Chinese sentences to be on par with English OIA through better integrated node type and edge type specific to Chinese OIA.
  - During the internship, my essay reading skills and coding skills improved substantially, also have research experience in OIE and logical reasoning.

Competitions

- **Kaggle — Feedback Prize - Predicting Effective Arguments Competition** — Jun 2022 - Aug 2022
  Team Leader — Sliver Medal (Top 2%)
  - Composed the baseline code and designed various data preprocessing strategies and model structures.
  - Using token classification instead of sequence classification in this competition. This new solution boosts our rank and saves a lot of time on training and inference. Acquire the ability to use the Transformers training API.
- **Kaggle — U.S. Patent Phrase to Phrase Matching** — Mar 2022 - Jun 2022
  Solo — Bronze Medal (Top 9%)
  - Composed the entire code pipeline with various combinations of pre-trained language models and structures.
  - Mastered multiple methods of fine-tuning such as mean pooling, multi-dropout, and layer normalization.
- **CCF Collegiate Computer System & Programming Contest** — Sep 2021 - Nov 2021
  Team Member — Top 5%
  - Utilized data science and machine learning pipeline to predict consumer default vs non-default.
  - Master the using of lightGBM and XGBoost models, and the model stacking method. Mastered real-world data science challenge with machine learning pipeline and team collaboration.

Projects

- **Rumor Prediction Model for Microblog Epidemic** — Mar 2021 - May 2021
  Solo Project
  - A Covid-19 rumor prediction model was built based on the pre-trained model BERT, which completes the deep learning model complete process.
  - In this project, I went through the whole process of idea generation, data search, reference papers, code practice, model building and article writing, and gained a very deep understanding of the NLP project process.
- **Emotional analysis case** — July 2021 - Sep 2021
  Group Leader
  - Sentiment analysis is performed on the movie review text dataset, and the bag-of-words model, N-gram model, loss function, and gradient descent function are implemented manually.
  - Through this project, I have gained a deeper understanding of the principles of deep learning, the real usage of gradient descent, and a deeper understanding of statistical-based language models.

Personal Summary

- **Languages**: Mandarin, English
- **Frameworks**: PyTorch, Scikit, NLTK