

Yunze Xiao

Doha, Qatar | yunzex@andrew.cmu.edu | [Personal Website](#) | +974 3996 8706

SUMMARY

- Natural Language Processing Enthusiast
- Adaptable to new environments and quick learner of new skills
- Language skills: English (high level), Chinese (native), Japanese (basic), Spanish(basic)

EDUCATION

Carnegie Mellon University **Class of 2026**

B.S. Computer Science

COURSES:

Natural Language Processing | Principles of Imperative Computation | Probability for Computer Scientists | Information System Milieux

PAST PROJECT

A Transformer-based Attention Flow Model for Intelligent Question and Answering Chatbot

- Improved the performance of the BiDaF model
- A [presented paper](#) in the 14th ICCRD

ONGOING PROJECT

[A dataset for Offensive Language in Chinese](#)

- Developing dataset for "abstract language" in Chinese
- Enhancing model's offensive phrase detection
- Expanding understanding of internet style
- Dataset size: 8218 instances of "abstract language" with subversive phrases.

Exploring the Current Approaches and Challenges in Chinese Offensive Language Identification and Detection

- Conducting a comprehensive survey on offensive language detection in Chinese.
- Examining and analyzed current benchmarks and approaches in the field.
- Investigating specific models and tools designed to address the unique challenges of detecting offensive language in Chinese.
- Identifying potential avenues for further research, considering the cultural and linguistic complexities of the Chinese language.

Exam Master: a LLM based practice generator

- Finetuning LLM to generate course-specific questions.
- Providing a tool to track student's progress and weakness.
- Incorporate personal trend into the model.
- Improving the model through learner-sourcing.
- Identifying how to system can help student improve their grades.

WORK EXPERIENCE

Curriculum Support Developer

- Creation of question banks for quizzes and exams. This includes reviewing quizzes and exams from past course iterations to extract questions. The developer will help create a database to organize all these questions by topic, assign labels, and add comments
- help develop a web application that can be used to browse the database. Responsibilities include writing and testing code, helping design the software architecture, debugging, and deploying the web application.

Course Assistant for 15-110

- Help students understand key concepts of 15-110
- Grade student's homework and provide feedback.

NLP Research Intern

- Working on real-life problem in the realm of Arabic NLP, such as stance classification and fake news detection.
- Participating WANLP 2023 shared tasks in propaganda detection and publishing a working note in achieve semi-SOTA solutions.

AWARD

CMU-Q Dean's List 2022 Fall

LEADERSHIP

Head of the Student Academic Committee

- Helped connect student with the University faculties.
- Organized the townhall event in Fall 2023.