

Chenxi Li

Email: cl91@uw.edu Phone: 206-409-6043

EDUCATION

- | | |
|--|------------------------------|
| University of Washington | Seattle, USA |
| • <i>Ph.D. in Computational Linguistics</i> | <i>Sep. 2024 – present</i> |
| <i>M.S. in Computational Linguistics; GPA: 3.98</i> | <i>Sep. 2022 – May 2024</i> |
| University of Pennsylvania | Philadelphia, USA |
| • <i>M.S.Ed. in Teaching English to Speakers of Other Languages; GPA: 3.88</i> | <i>Sep. 2014 – May 2016</i> |
| Beijing Language and Culture University | Beijing, China |
| • <i>B.A. Teaching Chinese to Speakers of Other Languages; GPA: 3.76</i> | <i>Sep. 2010 – Jun. 2014</i> |

PUBLICATION

- Z. Zerong, **C. Li**, X. Lin, J. Chen, F. Xia. 2025. A Systematic Survey of Claim Verification: Corpora, Systems, and Case Studies. Findings of the Association for Computational Linguistics: EMNLP 2025.
- T. Zhang, Z. Sun, S. Z, **C. Li**, N. Abernethy, B. Lam, F. Xia, M. Yetisgen. 2025: Thinking, Fine-Tuning, and Dictionary-Enhanced LLM Systems for Chemotherapy Timeline Extraction. In Proceedings of the 7th Clinical NLP Workshop (co-located with EMNLP 2025).
- **C. Li**, Y. Tian, Z. Zerong, Y. Song, F. Xia. 2024. *Challenging Large Language Models with New Tasks: A Study on their Adaptability and Robustness*. In Findings of the Association for Computational Linguistics: ACL 2024.

RESEARCH EXPERIENCE

- | | |
|-----------------------------------|----------------------------|
| Research Projects | Seattle, WA |
| • <i>University of Washington</i> | <i>Sep. 2023 - present</i> |
- **Chemotherapy Extraction from Clinical Notes:** Developed multiple LLM-based systems for extracting chemotherapy treatment timelines from clinical notes, including prompt engineering with thinking, dictionary-enhanced extraction using domain-specific knowledge, and supervised fine-tuning with direct preference optimization on Qwen3 models up to 14B parameters; our team won the leaderborad at ChemoTimelines 2025 Shared Task.
 - **Claim Verification Survey:** Conducted a survey of 198 claim verification papers summarizing recent advance in this area on both system design and corpus creation; used two case studies to demonstrate the problem with corpus quality, sentence selection in traditional pipeline, and claim decomposition.
 - **Course-scheduling System:** Collaborated in the design of solving course-schedule optimization problem for the linguistics department at UW by using integer linear programming. Translated real life concerns such as departmental requirements and instructor preferences into mathematical constraints to generate an optimized annual course schedule; implemented the code of this end-to-end system.
 - **LLMs Evaluation on Novel Tasks:** Developed and evaluated six novel tasks to assess LLMs' performance, creating synthetic datasets for automated testing and conducting human evaluations for comparison. Enhanced model performance through prompting strategies such as Chain of Thought techniques. Identified common errors in LLMs for models' architectural improvements.

- **Prompt Engineering for Chinese CLIP:** Enhanced the performance of Chinese CLIP in zero-shot image classification tasks on Caltech101 and ImageNet. Replaced single-word class labels (e.g., ‘dog’) with GPT-4 generated description (e.g., ‘dog is a furry animal...’) for text encoder during testing and incorporated bilingual embedding ensembling (English and Chinese) to optimize the results.
- **Emotion Detection on Friends Series:** Collaborated in the course project on sentiment analysis to classify transcripts from the TV series ‘Friends’ into emotion labels. Tested three approaches: probability lexicon, SVM, and adding a linear classifier to fine-tune BERT; Applied majority voting ensembling to enhance the accuracy of the model by combining individual predictions.
- **Multi-document Summarization for News Report:** Collaborated on the design and development of the text summarization task and evaluation pipeline. Used traditional TF-IDF method as baseline; Fine-tuned LLaMA2 using Low-Rank Adaptation (LoRA) to improve the model performance by 5%. Performed analysis on both the input and output data to understand the model performance.

WORK EXPERIENCE

- **Director** Shanghai, China
Walnut English Department, PalmDrive Jun. 2019 - Aug. 2022
 - Founded and led Walnut English Department, providing tailored standardized test prep courses to about 2000 college students annually
 - Oversaw all aspects of product marketing, development, implementation, and operation within the department
 - Collaborated with engineering teams to design and modify the customer management system, resulting in improved service quality and streamlined department operations
- **Senior Teacher** Shanghai, China
Higher Education Consulting, PalmDrive Jun. 2016 – Jun. 2019
 - Led the curriculum development of English standardized tests and implemented the curriculum within company’s technological frame
 - Taught GRE Quantitative and GMAT Quantitative, TOEFL and IELTS speaking and writing to online classes of 50+

SKILLS

- **Natural Language:** Mandarin (Native), English (Proficient)
- **Programming Language:** Python, C#
- **Toolkits:** PyTorch, scikit-learn, NLTK, spaCy, hugging face Transformers Numpy, PuLP(optimization), LATEX, Neo4j