Hainiu Xu

N5.08 Bush House, 30 Aldwych, London | hainiu.xu@kcl.ac.uk | https://seacowx.github.io

EDUCATION

King's College London   *London, UK*
PhD in Computer Science   *Oct. 2023 - June 2027 (Expected)*
- Supervisors: Professor Yulan He, Professor Caroline Catmur, Dr. Jinhua Du
- Studentship: EPSRC iCASE jointly funded by UKRI, Huawei London Research Centre, King's College London
- Project: *"Character-Centric Narrative Understanding"*

University of Pennsylvania   *Philadelphia, PA*
M.S.E in Data Science   *Aug. 2021- May 2023*
- *Outstanding Research Award*
- *Best Thesis Runner-up Award*
- Main course modules: Machine Learning; Deep Learning; Computational Linguistics; Principles of Deep Learning, Advanced Computational Linguistics
- GPA: 4.00 / 4.00

University of California, Davis   *Davis, CA*
B.S. in Statistics with Highest Honor (Statistical Data Science Track)   *Sept. 2016 - Jun. 2020*
- *Highest Honor (Summa Cum Laude)*
- Honor's Thesis: *"Application of Curve Registration Methods on Analyzing Wearable Device Data."*
- Supervisor: Professor Jane-Ling Wang
- Major GPA: 3.94 / 4.00 (overall GPA: 3.84 / 4.00)

THESIS

[1] "Fine-grained AND Coarse-grained Causal Reasoning in Procedural Texts", Master's Thesis, University of Pennsylvania, [LINK]

[2] "Application of Curve Registration Methods on Analyzing Wearable Device Data.", Undergraduate Thesis, University of California-Davis, [LINK]

RESEARCH EXPERIENCE

EPSRC-iCASE: Character-Centric Narrative Understanding   *London, UK*
UKRI, Huawei London Research Centre, King's College London
Supervisors: Prof. Yulan He, Prof. Caroline Catmur, Dr. Jinhua Du   *Oct. 2023 -*
- Develop Theory-of-Mind Reasoning benchmarks for LLMs.
- Construct euro-symbolic framework for enhancing LLMs' Theory-of-Mind reasoning capabilities.
- Investigate application of Theory-of-Mind reasoning in cognitive reappraisal.

Research Assistant: Procedural Reasoning   *Philadelphia, PA*
University of Pennsylvania, PennNLP
Supervisor: Prof. Chris Callison-Burch   *May. 2022 - Sept. 2023*
- Investigated methods for decomposing compound questions for procedural reasoning tasks under a few-shot setting.

- Conducted template-based question decomposition using GPT3 and T5.
- Created a metric for evaluating decomposition results for procedural reasoning tasks.
- Conducted image synthesis using procedural instructions with DALL·E-mini and CLIP.
- Investigated methods for composing logical relationship between multi-hop questions and corresponding zero-hop premises.
- Conducted error analysis on SOTA models in entity state tracking.
- Wrote conference paper as co-first author.

Independent Research: Visual-Guided Procedure Generation *Philadelphia, PA*
University of Pennsylvania, PennNLP
Supervisor: Prof. Chris Callison-Burch *May. 2022 - May. 2023*
- Built Crawler to gather parallel procedural data that contain texts and images.
- Investigated methods for goal inference based on the sequence of images.
- Investigated step-generation based on activity goal and semantics from images.

Honor's Thesis: Curve Registration on Wearable Device Data *Davis, CA*
UC Davis, Department of Statistics
Supervisor: Professor Jane-Ling Wang. *Jun. 2019 - Jun. 2020*
- Applied various time warping algorithms to functional data collected from wearable devices.
- Wrote a comprehensive literature review for the major warping methods.
- Compared time warping methods for wearable device data.
- Measured the effectiveness of warping methods with functional principal component analysis results on thewearable device data.

Research Training Group (RTG) Project: Analysis of Wearable Device Data *Davis, CA*
UC Davis Department of Statistics, National Science Foundation (NSF)
Supervisor: Professor Jane-Ling Wang *Dec. 2018 - Jun. 2019*
- Studied theories of functional data analysis including smoothing, clustering, and functional principal componentanalysis.
- Conducted exploratory functional data analysis with wearable device data.
- Extracted user activity patterns using functional principal component analysis.

WORK EXPERIENCE

Research Intern *London, UK*
Huawei London Research Center
Supervisor: Dr. Jinhua Du *Jun. 2024 - Sept. 2024*
- Worked on character relation extraction using LLMs.
- Built an event-centric, saliency-aware character relation extraction framework.
- Designed character saliency as well as character relationship saliency measure using graph centrality and PageRank.
- Achieved 82% precision score on long narratives.

AI Engineer Intern *Beijing, China*
Schlumberger BGC
Supervisors: Dr. Ping Zhang, Dr. Qing Liu, Dr. Peng Jin *Sept. 2020 - May. 2021*
- Conducted text mining on drilling reports written in Chinese.
- Conducted fine-tuning on Chinese word vectors for the oil and gas industry.
- Studied the behaviour of fine-tuning algorithms under a small-sample fine-tuning setting.
- Conducted a comprehensive evaluation of the performance of state-of-the-art language models (BERT, ALBERT,ELECTRA, etc.) on representing words from drilling reports.
- Built mono-lingual and multi-lingual classification models for classifying Chinese drilling reports.

• Deployed the classification model as a web application using Flask.

TEACHING

Teaching Assistant   *Philadelphia, PA*
University of Pennsylvania, CIS522 Deep Learning
Instructors: Prof. Lyle Unger, Prof. Konrad Kording   Jan. 2023 - May 2023
Give weekly lectures
Prepare lecture slides and tutorial solutions
Mentor course projects
Grade assignments and course projects

Teaching Assistant   *Philadelphia, PA*
University of Pennsylvania, CIS530 Computational Linguistics
Instructor: Prof. Mark Yatskar   *Aug. 2022 - Jan. 2023*
• Manage Piazza (course online QA platform)
• Make weekly quizzes
• Hold office hours
• Grade homework
• Mentor course projects

Tutor   *Davis, CA*
UC Davis: Academic Assistant and Tutoring Center   *Sep. 2018 - Dec. 2018*
• Making interactive questions regarding the material and making up example questions.
• Provided individual tutoring for peers from STA 131A: Introduction to Probability Theory.

EXTRACURRICULAR ACTIVITIES

UC Davis Symphony Orchestra   *Davis, CA*
Clarinet   *Jan.2018 - Jun. 2020*
• Rehearsed and performed in numerous concerts at the Mondavi Centre.
• Performed as the principal clarinet in various concerts.

UC Davis University Concert Band   *Davis, CA*
Principal Clarinet   *Jun. 2016 - Dec. 2018*
• Rehearsed and performed in numerous concerts at the Mondavi Centre.
• Led the clarinet section rehearsal.

UC Davis Chamber Music– Clarinet Trio   *Davis, CA*
Principal Clarinet   *Jan. 2019 - Mar. 2019*
• Rehearsed and performed in numerous concerts at the Pizter Centre.
• Conducted and directed the rehearsal of the chamber music group.

SKILLS & INTERESTS

Language: Fluent in Mandarin and English (GRE: 329 (Verbal Reasoning: 161, Quantitative Reasoning: 168))
Programming Languages: Python, R, C, MATLAB, Bash
Machine Learning Libraries: PyTorch, Huggingface Transformers, Numpy, Scikit-Learn, HyperOpt