Simon Yu

✉ simon011130@gmail.com
✉ yu.chi@northeastern.edu
⌨ @simonucl
in in/simonucl

Last updated on 2nd October 2024

EDUCATION

Northeastern University *Boston, MA*
PhD Student Sep 2024 - Now
- Research Area: *Conversational Agent, Alignment, Safety in ML*
- Supervisor: Weiyan Shi

The University of Edinburgh *Edinburgh, UK*
MSc. (by Research) Student Oct 2023 - Aug 2024
- Research Area: *Safety in NLP, Instruction Tuning*
- Supervisor: Dr. Jeff Z. Pan and Prof. Pasquale Minervini

The University Of Edinburgh *Edinburgh, UK*
BSc Artificial Intelligence and Computer Science Sep 2019 - Jul 2023
- First Honour (GPA 4.0/4.0)
- Supervised by Dr Jacques Fleuriot
- Related Courses: Natural Language Understanding, Generation, and Machine Translation, Text Technologies for Data Science, Software Engineering and Professional Practice

PUBLICATIONS

Language Model Merging in Iterative Preference Learning
Chenmien Tan, **Simon Yu**, Shay B. Cohen
Under review, 2024

Diversify and Conquer: Diversity-Centric Data Selection with Iterative Refinement
Simon Yu*, Liangyu Chen*, Sara Ahmadian, Marzieh Fadaee
Under review, 2024

Fine-Tuning Large Language Models with Sequential Instructions
Hanxu Hu*, **Simon Yu***, Pinzhen Chen*, Edoardo M. Ponti
Under review, 2024

Is It Good Data for Multilingual Instruction Tuning or Just Bad Multilingual Evaluation for Large Language Models?
Pinzhen Chen, **Simon Yu**, Zhicheng Guo, Barry Haddow
The 2024 Conference on Empirical Methods in Natural Language Processing (Main, EMNLP-2024)

Evaluating the Adversarial Robustness of Retrieval-Based In-Context Learning for Large Language Models
Simon Yu, Jie He, Pasquale Miniverini, and Jeff Pan
Conference on Language Modeling 2024 (COLM-2024)

Instances and Labels: Hierarchy-aware Joint Supervised Contrastive Learning for Hierarchical Multi-Label Text Classification
Simon Yu*, Jie He*, Victor Gutierrez-Basulto, and Jeff Pan
The 2023 Conference on Empirical Methods in Natural Language Processing (Findings, EMNLP-2023)

EXPERIENCE

Research Associate *Oct 2023 - Sept 2024*
Cohere for AI
- Research Topic: Data selection and clustering for instruction fine-tuning
- Collaborators: Liangyu Chen, Marzieh Fadaee, Sara Ahmadian

Research Assistant *Oct 2022 - Sept 2023*
Institute for Language, Cognition and Computation (ILCC), University of Edinburgh
- Research Areas: Supervised Contrastive Learning in NLP, Few-shot learning, Commonsense Question Answering
- 1 Paper Accepted by ACL 2023 Main (Poster)
- Supervisors: Dr Jeff Z. Pan

Machine Learning Engineer Intern   *June 2023 - Sept 2023*

Expedia Group, Inc.   London, UK

- Specialization in the Expedia For Business (E4B) teams, primarily focused on deploying AI models more efficiently for business applications.
- Designed and assisted in the deployment of Machine Learning models through a more streamlined pipeline using ONNX and CI/CD tools. This included migrating the deployment pipeline from Jenkins to GitHub Actions to improve CI/CD processes.
- Employed CI/CD tools (Spinnaker, Github Actions) and Cloud Services (AWS S3 Bucket, Lambda)

Research Intern   *March 2022 - Sept 2022*

Huawei Noah's Lab R&D   Edinburgh, UK

- Specialization in E-commerce Knowledge Graphs, Entity Linking, and Disambiguation from unstructured data
- Enhanced the performance of the entity extraction module, and contributed to the benchmarking work, fostering better integration between unstructured data and the product's knowledge graph.
- Assisted in the design and development of the product's Knowledge Graph. Leveraged data processing tools such as Apache Spark and Hadoop for product data processing. Employed CI/CD tools (Nexus Repository, Jenkins).

Honours & Awards

Outstanding Dissertation Award   *2023*

Scored top (90%) in the final year dissertation among the class

Conference Reviewer *EMNLP-2023, NAACL-2024*