Yanchen Liu

Graduate School of Arts and Sciences
Harvard University
Cambridge, MA, USA, 02138

+1 6173970175
yanchenliu@g.harvard.edu
liuyanchen1015.github.io

EDUCATION

Harvard University *2022 - 2024*
MS in Data Science
Cross-Registration in Computer Science at **MIT**

Advisors: Prof. Jiaqi Ma and Prof. Himabindu Lakkaraju

Technical University of Munich *2018 - 2022*
BS in Computer Science with Highest Honors
Minor in Computational Linguistics at **Ludwig Maximilian University**

Advisors: Timo Schick and Prof. Hinrich Schütze
Major GPA: 1.2/1.0 (3.97/4.0) Minor GPA: 1.0/1.0 (4.0/4.0)
Rank: **top 1%** with most courses passed with full scores (1.0/A+), particularly in all math

PUBLICATIONS

[1] PrivacyLens: Evaluating Privacy Norm Awareness of Language Models in Action
Yijia Shao, Tianshi Li, Weiyan Shi, **Yanchen Liu**, Diyi Yang
In Proceedings of the 38th Annual Conference on Neural Information Processing Systems (NeurIPS 2024)

[2] Decoding Susceptibility: Modeling Misbelief to Misinformation Through a Computational Approach
Yanchen Liu, Mingyu Derek Ma, Wenna Qin, Azure Zhou, Jiaao Chen, Weiyan Shi, Wei Wang, Diyi Yang
In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP 2024)

[3] Confronting LLMs with Traditional ML: Rethinking the Fairness of Large Language Models in Tabular Classification
Yanchen Liu, Srishti Gautam, Jiaqi Ma, Himabindu Lakkaraju
In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL 2024)

[4] DADA: Dialect Adaptation via Dynamic Aggregation of Linguistic Rules
Yanchen Liu, William Held, Diyi Yang
In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)

[5] Task-Agnostic Low-Rank Adapters for Unseen English Dialects
Zedian Xiao, William Held, **Yanchen Liu**, Diyi Yang
In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)

[6] SMoA: Sparse Mixture of Adapters to Mitigate Multiple Dataset Biases
Yanchen Liu*, Jing Yan*, Yan Chen*, Jing Liu, Hua Wu
In ACL Workshop on Trustworthy Natural Language Processing, 2022

[7] Semantic-Oriented Unlabeled Priming for Large-Scale Language Models
Yanchen Liu, Timo Schick, Hinrich Schütze
In ACL Workshop on Simple & Efficient Natural Language Processing, 2022
Oral Presentation

ACHIEVEMENTS

best.in.tum — *Apr. 2020*
promotion of the best students — *TU Munich, DE*

RESEARCH EXPERIENCE

Stanford NLP Group — *Oct. 2022 - Present*
Visiting Research Assistant — *Palo Alto, CA*
Advisor: Prof. Diyi Yang

- ❏ Proposed Dialect Adaptation via Dynamic Aggregation (DADA), a compositional and modular approach to enhance the dialectal robustness of models trained on Standard American English across multiple dialects simultaneously, from a finer-grained perspective to accommodate dialect flexibility [4].
- ❏ Introduced HyperLoRA, a scalable, task-agnostic method that incorporates expert linguistic knowledge to enable resource-efficient dialect adaptation through the use of hypernetworks to disentangle dialect-specific and cross-dialectal information [5].
- ❏ Formulated a computational approach to model users' susceptibility to misinformation based on their online activities, using observable sharing behavior as a proxy, and enabling large-scale analysis of its correlation with social and psychological factors [2].

Harvard AI4LIFE Group — *Mar. 2023 - Present*
Research Assistant — *Cambridge, MA*
Advisor: Prof. Himabindu Lakkaraju

- ❏ Analyzed how LLMs exhibit inherent social biases inherited from their pre-training corpora, and investigated the fairness implications of LLMs when making predictions on tabular data, in comparison with traditional machine learning models [3].

LMU Center for Information & Language Processing — *Jun. 2021 - Nov. 2021*
Research Assistant — *Munich, DE*
Advisor: Prof. Hinrich Schütze

- ❏ Proposed Semantic-Oriented Unlabeled Priming (Soup), a novel approach by retrieving and leveraging semantically similar unlabeled examples for enhancing the few-shot performance of pre-trained LMs. And introduced bag-of-contexts priming, a new priming strategy that is more suitable for this setting and enables the usage of more examples than fit into the context window [7].

TALKS

Stanford NLP Talk — *Nov. 2023*
Dynamic Aggregation and Auto-Discovery of Linguistic Features

Stanford NLP Lightning Talk — *Oct. 2023*
LLM for More Research: Empowering Linguistic and CSS Research with LLMs

PROFESSIONAL SERVICE

Mentor: Rodrigo Nieto (BS/MS@Stanford, Sep. 2023 - Mar. 2024), Azure Zhou (BS@Stanford, Jun. 2023 - Sep. 2023), Mary Williamson (MS@Stanford, Jun. 2023 - Sep. 2023)

Reviewer: ARR 2024, COLM 2024, NAACL 2024 SRW

SKILLS

Programming Languages: C/C++, Java, Python, OCaml, Verilog, MIPS Assembly, SQL...
Language Proficiency: English - TOEFL 111, German - DSH2, Chinese - Native
Also hobbies: Soccer, Go (3 Dan)