Leo Zeyu Liu

206-876-0514 | zeyuliu2@cs.washington.edu | **leo-liuzy**

Education

University of Washington *Seattle, WA*

BS/MS in Computer Science *Sept. 2017 – June 2022 (Expected)*

BS GPA: 3.88/4.0 *MS GPA: 3.95/4.0*

Advisor: Professor Noah A. Smith and Professor Shane Steinert-Threlkeld

Selected Coursework

Machine Learning, Natural Language Processing, Artificial Intelligence, Statistical Learning, Interactive Learning, Information Theory I, Deep Learning, Algorithms, Data Structure, Database Management, Software Engineering, Operating System, Programming Languages, Syntax, Semantics, Linear Optimization

Experience

AI2 *Jan. 2021 — Mar. 2021*

Software Engineer managed by Evan Pete Walsh *Seattle, WA*

- Adding recent influence function-based methods to AllenNLP Interpret

Noah's Ark *Jan. 2019 — Present*

Research Assistant advised by Professor Noah A. Smith *University of Washington, Seattle, WA*

Probing Across Time [1] *May 2020 — Present*

- Collaborator(s): Yizhong Wang, Jungo Kasai
- Raise a new dimension to understand the largely unknown pretraining process of language model.
- Find out different learning patterns of different kinds of knowledge.

Lifelong Learning in NLP *Jan. 2020 — May 2020*

- Collaborator(s): Yizhong Wang
- Re-implement and attempt to improve a simple sampling-based lifelong learning algorithm — MBPA++

BERT Representation Analysis *Sept. 2019 — Dec. 2019*

- Collaborator(s): Yizhong Wang
- Use bottleneck network to find out that decomposed BERT representation suffice to perform well on SQuaD v1.1 in very low dimension space.
- LayerNorm is expressive on its own; simply training LayerNorm of pretrained BERT will contribute over 70% to the performance on SQuaD v1.1.

Neural Persona *Jan. 2019 — Sept. 2019*

- Collaborator(s): Dr. Dallas Card
- Design learning-based system by using Variational Auto-Encoder to extend neural topic modeling to also model entities in purely raw text corpus.

Computation, Language, and Meaning Band of Researchers *Jan. 2019 — Present*

Research Assistant advised by Professor Shane Steinert-Threlkeld *University of Washington, Seattle, WA*

Auto-generate contrast set [2]

- Collaborator(s): Chuanrong Li, Shengshuo Lin, Xinyi Wu, Xuhui Zhou
- Based on the ACE parser, a well-known engineered grammar parser, to build pipeline to automatically generate contrast set. For example, changing a sentence to its passive form should not affect its label, yet adding modality (e.g., "may") to it will.
- We test the viability of this idea on Natural Language Inference task, but our pipeline is extensible to other tasks with different heuristic design.

Research Computing Club *Sept. 2018 — Sept. 2019*

Undergraduate Liaison *University of Washington, Seattle, WA*

- Design and lead weekly hands-on workshops to help interested members to start on the Kaggle competition.
- Mentor members in the Club to use UW High Performance Computing Cluster.

BioASQ Challenge *April. 2018 — June. 2019*
Research Assistant *University of Washington, Seattle, WA*

- Collaborator(s): Will Kearns, Professor Xiaosong Li
- Design the Answer Generation, data storage, preprocessing component.
- Pre-training Question Answering (QA) model on CNN/Daily Mail dataset, and finetune on QA dataset in medical domain.

Publications

* = equal contribution, α = sorted alphabetically

[1] **Leo Z. Liu***, Yizhong Wang*, Jungo Kasai, Hannaneh Hajishirzi, and Noah A. Smith. Probing across time: What does roberta know and when? In ***submission to NAACL***, 2021

[2] Chuanrong Li$^{\alpha}$, Lin Shengshuo$^{\alpha}$, **Leo Z. Liu**$^{\alpha}$, Xinyi Wu$^{\alpha}$, Xuhui Zhou$^{\alpha}$, and Shane Steinert-Threlkeld. Linguistically-informed transformations (LIT): A method for automatically generating contrast sets. In *Proc. of EMNLP BlackboxNLP Workshop (Poster)*, 2020. arXiv: 2010.08580

Honors & Rewards

Dean's List 2017 – 2020
Citadel UW Datathon 2nd Place ($2500) 2018

Professional Service

Student Volunteer:

- Annual Meeting of the Association for Computational Linguistics (ACL): 2020
- Conference on Empirical Methods in Natural Language Processing (EMNLP): 2020

Teaching Experience

CSE 446: Machine Learning, Undergraduate TA *April 2019 – June 2019*
with Professor Kevin Jamieson and Professor Anna R. Karlin *Seattle, WA*

- Assisted with course planning and development, led a weekly discussion section, and held weekly office hours. Developed a new section material for linear algebra in Machine Learning context.

CSE 446: Machine Learning, Undergraduate TA *Sept. 2019 – Dec. 2019*
with Professor Sewoong Oh *Seattle, WA*

- Assisted with course planning and development, led a weekly discussion section, and held weekly office hours. Developed a new section material for Maximum Likelihood Estimation.

CSE 546: Machine Learning, Graduate TA *April 2020 – June 2020*
with Professor Kevin Jamieson and Professor Jamie Morgenstern *Seattle, WA*

- Assisted with course planning and development, led a weekly discussion section, and held weekly office hours.

CSE 546: Machine Learning, Graduate TA *Sept. 2020 – Dec. 2020*
with Professor Kevin Jamieson and Professor Jamie Morgenstern *Seattle, WA*

- Assisted with course planning and development, held weekly office hours, attend online Question Answering. Developed a new section material for Stein's paradox and a new topic modeling assignment

Technical Skills

Programing: Python, Java, C/C++, SQL, LISP-like, Bash, LaTeX
Frameworks/Libraries: PyTorch, `scikit-learn`, AllenNLP, Numpy
Language: Chinese (Native), English (Proficient)

Last Updated on 31st January 2021