Yifei Li

Master Student · University of Pennsylvania

liyifei@seas.upenn.edu | realliyifei.github.io | realliyifei | realliyifei | @realliyifei

*I am **open to any research topic related to AI**, as long as I can develop my own research skills and taste during my Ph.D. Currently, ...*
*My **research agenda** is motivated by two overarching, complementary goals: to advance intelligence in NLP and multimodal AI to understand concepts in a more composite, consistent, and hierarchical way; and to ground these models to serve human activities more robustly and fairer*
*My **domains** are: Natural Language Processing, Multimodal Learning, Neuro-Symbolic Approaches, Human-Centric AI and Fairness*

Education

University of Pennsylvania (UPenn) — *Philadelphia, PA, USA*

M.S.E in Data Science — *GPA: 3.97/4.0 | 2023**

- *Courses*: Computational Linguistics, Advanced Computer Vision, Interactive Fiction and Text Generation, Machine Learning, Theory of Machine Learning, Deep Learning, Convex Optimization, Learning in Robotics, Statistics in Data Science, Big Data Analysis

Oklahoma State University (OSU) — *Stillwater, OK, USA*

B.S. in Computer Science, B.S. in Mathematics, and B.S.B.A. in Management — *GPA: 3.8/4.0 | 2019*

- *Courses*: Advanced Linear Algebra, Mathematical Analysis, Probability and Mathematical Statistics, Computer Networks, Database

Sun Yat-Sen University (SYSU) — *Guangzhou, China*

B.M. in Management *(Notes: International joint-degree program associated with OSU)* — *GPA: 3.8/4.0 | 2019*

Publications

Papers / Manuscripts [2]

Gender Bias on Self-Supervised Learning
Yukai Yang*, Yifei Li*, Vicente Ordonez, Mark Yatskar
Working.
URL: `https://to/be/release.com` 2023

Conceptor-Aided Debiasing of Contextualized Embeddings
Yifei Li, Lyle Ungar, João Sedoc
EMNLP 2022 Rejected (Review Score 3.5/3.5/3); Resubmitting to another NLP conference.
URL: `https://arxiv.org/abs/2211.11087` 2022

Skills

Coding Languages	Python, Java, MATLAB, R, C++, SQL, JavaScript, CSS, PHP
Libraries and Tools	PyTorch, Hugging Face, Faiss, Spark, Shell, Git, LaTeX
AI Models	BERT family, GPT family, CLIP, Stable Diffusion, MoCo, YOLO, SOLO, GAN family

Research Projects

Conceptor-Aided Debiasing of Contextualized Embeddings — *UPenn*

Advised by Prof. João Sedoc and Prof. Lyle Ungar - Paper — *Jan. 2022 - Jun. 2022*

- Use conceptors–a soft projection method–to identify and remove the bias subspace in contextual embeddings in BERT and GPT and reach SOTA performance. Two methods of applying conceptors are proposed: (1) bias subspace projection by post-processing; and (2) a new architecture, conceptor-intervened BERT (CI-BERT), which explicitly incorporates the conceptor projection into all layers during continued training.
- Show the importance of carefully constructing the bias subspace. The best results are obtained by removing outliers from the list of biased words, intersecting them (using the conceptor AND operation), and computing their embeddings using the sentences from a cleaner corpus.

Gender Bias on Self-Supervised Learning — *UPenn*

Advised by Prof. Mark Yatskar and Prof. Vicente Ordóñez — *Aug. 2022 - Present*

- Find that models even pretrained on non-human images would lead to gender bias after fine-tuning, due to the implicit confounders.
- Try to debias by finding and removing the biased training images by (1) tracing the gradient between model pretraining and gender classifier finetinuning following the TracIn method (2) computing the nearest neighbor similarity of embeddings.

Probing CLIP Zero-Shot Ability — *UPenn*

Advised by Prof. Mark Yatskar (Independent Study) — *Jan. 2022 - May. 2022*

- Test and evaluate the zero-shot ability of CLIP model, figure out that CLIP performs poor on fine-grained dataset e.g. iNaturalist.
- Try to build a language utility to help users ask the right questions to CLIP: leverages the sentence-BERT to cluster different types of descriptions from the web, then compares and ranks their CLIP similarity scores.

Improving Consistency of Vision-Language Multimodality *UPenn*

Advised by Prof. Chris Callison-Burch and Prof. Mark Yatskar (Master Thesis) *Aug. 2022 - Present*

- Try to improve consistency of stable diffusion by (1) fine-tuning on datasets CLEVR and Visual Genome with the help of counterfactual data augmentation (2) combining the concepts caught by textual inversion and grounded by GPT3 to render consistent scenes in textual game.

Neuro-Symbolic Dual-System on Task-Oriented Dialogue Generation *UPenn*

Advised by Prof. Chris Callison-Burch and Dr. Lara Martin (Research Course Project) - Report *Mar. 2022 - May. 2022*

- Adapt novel neuro-symbolic dual-system to improve the consistency and coherence in task-oriented dialog generation: build a user belief states to ground human knowledge and domain-specific constraints, then fine-tune GPT3 to generate utterance and verify the consistency via belief states by a symbolic parser, repeat this process until it is consistent then update the belief states. This model outperforms the single neural system in inform and success rate by a large margin.

Professional Activities

2022	**EMNLP,** Reviewer	
2022 Fa	**CIS 5300 Computational Linguistics,** Teaching Assistant (Create assignment; supervise course projects)	*UPenn*
2022 Sp	**CIS 522 Deep Learning,** Teaching Assistant (Lead recitation pods - notes; supervise course projects)	*UPenn*
2021 Fa	**CIS 520 Machine Learning,** Teaching Assistant (Lead recitation pods - notes; supervise course projects)	*UPenn*

Awards & Honors

2016-19	**President's Honor Roll,** Maintain Excellent GPA	*OSU*
2018	**Emeritus Math Faculty Scholarship,** Mathematics Department, 1-2 Student(s) Each Year	*OSU*
	SST Scholarship, Computer Science Department	*OSU*
2016	**Transfer Out-Of-State Achievement,** Top 15% GPA	*OSU*

Course Projects

Two-Stage Summarization with Pre-Trained Transformers *UPenn*

Coursework - GitHub, Report *2021*

- Present a two-stage summarization model using pretrained transformers, concatenatingan extractive model BERTSumEXT and an abstractive model, GPT2 or BART. By feeding the abstractor the extracted key information, this method can mitigate the disadvantages of both approaches and make the summaries more readable. It also reduces the need to truncate sentences for abstractive model due to the maximum token limit.

NLP and Text-to-Image Generation for Gameplaying: Steins;Gate *UPenn*

Coursework (funny) - GitHub *2022*

- Build an interactive textual game powered by NLP and text-to-image GAN where the storyline would change based on the procedures controlled by the player. Here, the storyline is represented as graph, the interactive texts are generated by GPT3, and the cutscenes are rendered by pixray.

SOLO and GRU for Hemostatic Plug Segmentation *UPenn*

Coursework - GitHub, Report *2021*

- Segment the hemostatic plug instance in 3D biomedical images: modify the SOLO model (Segmenting Objects by Locations) with a customized ResNet50 backbone for binary classification, with an addition of GRU for Feature Pyramid Network output to encode the sequence of images.

Product Match by Deep Learning in Computer Vision and Natural Language Processing *UPenn*

Coursework - GitHub, Report *2021*

- Match the same e-commerce products in 70k test dataset by exploiting their titles (BM25, Doc2Vec, BRET, Faiss) and images (CNN, DNN, VGG19, ResNet152) in 32k training entities with the help of zero-shot learning, triplet loss function, and KNN embedding, reaching 97% accuracy.

Hotel Cancellation Prediction Using 10+ Machine Learning Approaches *UPenn*

Coursework - GitHub, Report *2020*

- Use 10+ machine learning models (e.g. AdaBoost, XGBoost, SVM, and an ensemble method combining tuned neural network, tuned random forest, and decision tree) with SMOTE rebalance technique to predict the hotel booking cancellation, leading to 97.52% accuracy.

Inverse Reinforcement Learning on Gridworld *UPenn*

Coursework - GitHub, Report *2021*

- Compare and explore the reinforcement learning and inverse reinforcement learning models on different gridworld environments.

Miscellaneous

- Yifei has great interest in AI Art as hobby, especially text-to-image generation via diffusion models
- Yifei is a member of PennNLP community and Penn Data Science Group
- Yifei worked as analyst or technology consultant intern in several financial and accounting companies when majored in management, he fell in love with the emergence of intelligence and dramatically switched his interest to AI academia in the middle of his master at UPenn