Xiaoyu Yang

xiaoyu.yang94@hotmail.com Toronto, ON, Canada (+1) 613-531-1794
github.com/arielsho https://arielsho.github.io/ Google Scholar LinkedIn

Targeting a applied science full-time position.

RESEARCH INTERESTS & SKILLS

Natural Language Processing, Deep Learning, Machine Learning.
Python, PyTorch, Tensorflow, PyTorch Lightning, Scikit-learn, FAISS, PySpark, Docker, Gradio.

EDUCATION

Queen's University, Kingston, Canada *Sept. 2018 - Sept. 2022*
Doctor of Philosophy, Department of Electrical and Computer Engineering & Ingenuity Labs

University of Chinese Academy of Sciences, Beijing, China *Sept. 2016 - Jun. 2018*
Master, Institute of Information Engineering

Nankai University, Tianjin, China *Sept. 2012 - Jun. 2016*
Bachelor, Department of Computer and Control Engineering, GPA: 87.5/100 (top 10%)

EXPERIENCE

Software Engineer (ML/NLP) *Jul. 2023 – present*
Cresta AI, Toronto, Canada

AI Research Scientist *Oct. 2022 – Jun. 2023*
LG Toronto AI Lab, Toronto, Canada

- Developed an in-house retrieval-augmented large language model (RA-LLM) based on Fusion-in-Decoder with knowledge distillation, and evaluated it on large-scale question-answering datasets TriviaQA and NQ.
- Enabled distributed training with PyTorch DDP/Lightning, maximizing training efficiency and scalability.
- Reduced retrieval latency $<$ 500ms via index product quantization and FAISS GPU acceleration.
- Developed a chatbot utilizing GPT-3.5-turbo and LangChain to address user inquiries about product manuals.

Machine Learning Research Intern *Jun. 2022 – Aug. 2022*
Microsoft Research Lab (MSR), Montréal, Canada

- Created toy datasets to investigate the behavior of models on compositional splits and assessed the effectiveness of the attention regularization method.
- Proposed to disentangle the syntax and semantics modeling to improve compositional generalization in semantic parsing, and demonstrated its effectiveness through preliminary experiments on SMCalFlow-cs.

Machine Learning Research Intern *Jun. 2021 – Aug. 2021*
Samsung Research America (SRA), Mountain View, US

- Incorporated additional knowledge-aware supervision to improve neuro-symbolic visual concept learning.
- Improved both accuracy and training efficiency of the neuro-symbolic model on the VQA task.

Machine Learning Engineer Intern *Jun. 2018 – Aug. 2018*
Alibaba Damo Academy, Alibaba Group, Hangzhou, China

- Developed a neural information extraction model to extract client information from insurance policy documents.
- Enhanced the base model with attention mechanism and new training objective, improving accuracy by 4%.

PUBLICATIONS

Neuro-symbolic Natural Logic with Introspective Revision for Natural Language Inference ***TACL-2022***
Yufei Feng, Xiaoyu Yang*, Xiaodan Zhu, Michael Greenspan (*equal contribution)*

- Designed a neuro-symbolic framework with built-in interpretability that integrates long-studied natural logic with reinforcement learning for natural language inference.
- Proposed an introspective revision algorithm that leverages external knowledge to modify intermediate symbolic reasoning steps to alleviate spurious reasoning and improve training efficiency.

Exploring Decomposition for Table-based Fact Verification ***Findings of EMNLP-2021***
Xiaoyu Yang, Xiaodan Zhu

- Explored to better verify the complex statements against evidence tables by decomposing them into several simpler sub-problems. Achieved new state-of-the-art performance on table fact verification dataset TABFACT.
- Designed a program-guided approach to construct a pseudo decomposition dataset for decomposition model training, and further proposed a neural model to fuse sub-problems and their answers for verification.

Program Enhanced Fact Verification with Verbalization and Graph Attention Network ***EMNLP-2020***
Xiaoyu Yang, Feng Nie, Yufei Feng, Quan Liu, Zhigang Chen, Xiaodan Zhu

- Proposed a verbalization method to accumulate symbolic evidence inherently embedded in executable programs, and adapted margin loss to improve the performance of program selection module.
- Designed a graph attention network to combine the original evidence and verbalized operations for fact verification. Achieved state-of-the-art performance on a large-scale table fact verification dataset TABFACT.

SemEval-2020 Task 5: Counterfactual Recognition [Task Organizer] ***COLING-2020 SemEval***
Xiaoyu Yang, Stephen Obadinma, Huasha Zhao, Qiong Zhang, Stan Matwin, Xiaodan Zhu

- Organized the SemEval-2020 Task 5: counterfactual recognition. Scraped around 120k relevant sentences from web, obtained 20k annotated counterfactual samples using Mechanical Turk.
- Implemented the baseline models, organized the system evaluation and the workshop paper review process.

Enhancing Unsupervised Pretraining with External Knowledge for Natural Language Inference
Xiaoyu Yang, Xiaodan Zhu, Huasha Zhao, Qiong Zhang, Yufei Feng ***CanadianAI-2019***

- Integrated WordNet/ConceptNet knowledge into pre-trained neural networks for natural language inference.

Learning to Retrieve Entity-Aware Knowledge and Generate Responses with Copy Mechanism for Task-Oriented Dialogue Systems
C Tan, Xiaoyu Yang, Z Zheng, T Li, Y Feng, J Gu, Q Liu, D Liu, Z Ling, X Zhu ***AAAI-2021 DSTC-9***

- Worked in a team and made a system that first selects unstructured knowledge then generates responses.
- Designed a Retrieve & Rank model for knowledge selection that outperforms baseline by 35%.
- Conversation modeling with unstructured knowledge access, 2nd place in DSTC-9 Track-1 objective evaluation.

HONORS & AWARDS

- Borealis AI Fellowship Award (merit-based) [Link] *2020 - 2021*
- Vector Institute Scholarship (merit-based) [Link] *2018 - 2019*
- Second-grade scholarship (top 10%) in Nankai University (3 consecutive academic years) *2012 - 2015*

ACADEMIC SERVICES

- 2020 SemEval
- 2022 ARR