Yijun LIANG

yliang17@umd.edu +1 (240)-481-5811

EDUCATION

University of Maryland, College Park **Maryland, U.S.**
Department of Computer Science (P.h.D in Computer Science) Aug 2023-Present
University of California, Berkeley **California, U.S.**
Department of Industrial Engineering and Operations Research Department. (M.Eng in IEOR) Aug 2019-Dec 2020
Tsinghua University **Beijing, China**
Department of Industrial Engineering (B.E. in Operations Research) Aug 2015-Jul 2019
School of Economics and Management (Double Major, B.E. in Economics) Sep 2016-Jul 2019

PUBLICATIONS

Compute Like Humans: Interpretable Step-by-step Symbolic Computation with Deep Neural Network. **KDD, 2022**
*Peng, Shuai, Di Fu, Yong Cao, **Yijun Liang**, Gu Xu, Liangcai Gao, and Zhi Tang. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp. 1348-1357. 2022.*
Geodrl: A self-learning framework for geometry problem solving using reinforcement learning in deductive reasoning **ACL, 2023**
*Peng, Shuai, Di Fu, **Yijun Liang**, Liangcai Gao, and Zhi Tang. In Findings of the Association for Computational Linguistics: ACL 2023 (pp. 13468-13480).*

WORK IN PROGRESS

University of Maryland, College Park, Department of Computer Science, PhD I **Maryland, U.S.**
Advisor: Dr. Tianyi Zhou, Assistant Professor, Dept. Computer Science.
Diffusion Curriculum: Synthetic-to-Real Data Curriculum via Image-Guided Diffusion Nov 2023-May2024
- Used ***Diffusion models*** to generate controllable and diverse images tailored to challenging datasets with low quality.
- Develop a curriculum learning approach that utilizes generated data to enhance image classification performance with a CLIP-based ***Vision-Language Model***.

Bi-Directional Text-Image Generative Model Designing Utilizing Diffusion Architecture Nov 2023-Present
- Designed a ***bidirectional text-image generative model*** utilizing the framework of Diffusion model, facilitating a seamless text-to-image-to-text generation process within a singular integrated model.

PROJECT EXPERIENCE

ByteDance – Machine Learning Engineer **Beijing, China**
Advanced Models Design for Mathematics Symbolic Computation Mar 2021-Mar 2022
- Using ***BERT text transformer*** to select feasible fundamental computation transformations (FCTs) for a math formula accurately.
- Designed ***Huffman position embedding*** and ***attention matrix*** for a ***new Transformer encoder*** to take operator trees as input.
- Designed a new ***sub-tree pretraining task*** for FCTs selection. Taking all feasible FCTs for each individual subtree in the original operator tree as labels to enhance the learning of a ***global and local integrated position embedding***.
- While selecting 5 FCTs from 160+ candidates, the algorithm reaches ***a recall of 99.998% and precision of 74.6%.***

Text Classification and Generation with Language Models Mar 2022-Jun 2022
- Developed a ***DeBERTa model*** to clarify if an answer for a question is included in the video, with an accuracy of 86%.
- Established a ***DeBERTa encoder-decoder model*** with nucleus sampling to generate diversified question given multiple video texts.
- With trained model, over 100 million video-based Q&A pairs with high quality are generated as the content for the community.

Video Violation Detection with Vision-Language Models Sep 2022-July 2023
- Developed and optimized ***multi-modal models*** incorporating image and text encoders, alongside a designed fusion block, for the purpose of classifying video violations. Conducted experiments with a variety of image encoders ***(ResNet50, ViT, and Swin-T)*** in combination with text encoders ***(RoBERTa, BERT),*** enhancing the fusion block to attain superior performance outcomes.
- Designed and validated the viability of a new video moderation system (Policy Aware). Established and refined 12 violation categories (domain) within the system. Implemented and optimized 12 domain models to enhance the detection of violated videos.

RESEARCH EXPERIENCE

Vehicle Routing Optimization Based on ADP Algorithm **Sep 2018-Jun 2019**
Advisor: Dr. Tianhu Deng, Associate Professor, Dept. IE, Tsinghua University. Beijing, China
- Designed *an* ***approximate dynamic programming algorithm*** with an autoregressive value function to optimize this delivery problem with high dimensions of parameters and constrains.
- Approximated the value function through ***value function iterations*** algorithm with ***epsilon-greedy exploration.***
- Resolved the company's bottleneck by decrease the time-out times by 36% on average in simulations.

Theoretical Analysis and Algorithm Optimization for Rank Centrality Algorithm **Jul 2018-Aug 2018**
Advisor: Dr. Xiaodong Li, Associate Professor, Dept. Statistic, University of California, Davis. California, U.S.
- ***Complemented the theoretical deduction*** about the combination of ***Markov Chain*** and pairwise ranking for the literature.
- Found out the deterministic parameters for the algorithm convergence. With optimized parameters, the speed of convergence increased by 45% on average in simulations.