Patrick (Pengcheng) Jiang

Tel: +1-(217)2986575 | Email: pj20@illinois.edu | Homepage: pat-jj.github.io | GitHub: github.com/pat-jj

Education

University of Illinois at Urbana-Champaign — Champaign, USA
M.S. in Computer Science (Thesis-based) — Sep. 2022 – May 2024 (expected)

- GPA: 4.00/4.00 (Rank: NA); worked with Prof. Jimeng Sun and Prof. Jiawei Han
- Research interests: knowledge graphs, text mining, and machine learning for healthcare

Waseda University — Tokyo, Japan
B.E. in Computer Science and Engineering — Sep. 2017 – Sep. 2021

- GPA: 3.85/4.00 (WES); worked with Prof. Shinichi Honiden and Prof. Kenji Tei
- Research interests: formal methods, model checking, and machine learning

Publications

Graduate Research (Sep. 2022 ∼)

- **Pengcheng Jiang**, Cao Xiao, Tianfan Fu, Parminder Bhatia, Jimeng Sun, "Bi-level Contrastive Learning for Knowledge-Enhanced Molecule Representations", (under review at *KDD*'24). [pdf] [code]
- **Pengcheng Jiang**, Cao Xiao, Zifeng Wang, Parminder Bhatia, Jimeng Sun, and Jiawei Han, "TriSum: Learning Summarization Ability from Large Language Models", (commited to *NAACL*'24). [pdf]
- **Pengcheng Jiang**, Jiacheng Lin, Zifeng Wang, Jimeng Sun, and Jiawei Han, "GenRES: Rethinking Evaluation for Generative RE in the Era of LLMs", (commited to *NAACL*'24). [pdf]
- **Pengcheng Jiang**, Megan Amber Lim, Adam Cross, and Jimeng Sun, "MedKG: Empowering Medical Education with Interactive Construction and Visualization of Knowledge Graphs via Large Language Models", (under review at *JAMIA*). [pdf] [code]
- **Pengcheng Jiang**, Cao Xiao, Adam Cross, and Jimeng Sun, "GraphCare: Enhancing Healthcare Predictions with Personalized Knowledge Graphs", *ICLR*'24. [pdf] [code]
- Chaoqi Yang*, Zhengbang Wu*, **Patrick Jiang**, Zhen Lin, Junyi Gao, Benjamin Danek, and Jimeng Sun, "PyHealth: A Deep Learning Toolkit for Healthcare Predictive Modeling", *KDD*'23. [pdf] [doc] [code]
- **Pengcheng Jiang**, Shivam Agarwal, Bowen Jin, Xuan Wang, Jimeng Sun, and Jiawei Han, "Text-Augmented Open Knowledge Graph Completion via Pre-Trained Language Models", *ACL*'23. [pdf] [code]

Undergraduate Research

- **Pengcheng Jiang**, Kenji Tei, "OACAL: Finding Module-consistent Specifications to Secure Systems from Weakened User Obligations", *IEEE SSCI*'21. [pdf]
- **Pengcheng Jiang**, "CNN-based Diagnosis System on Skin Cancer using Ensemble Method Weighted by Cubic Precision" tech report. [pdf]

Academic Experience

University of Illinois at Urbana-Champaign — Champaign, USA
Research Assistant to Prof. Jimeng Sun — Sep. 2022 – May 2024 (expected)

- Developed PyHealth, a comprehensive deep learning toolkit designed for healthcare. [link] [code]
- Developed GraphCare, an innovative framework that constructs and uses patient's personalized KGs for enhanced clinical predictions.

- Developed GODE, a contrastive learning-based approach that synergizes molecule-centric biochemical knowledge graphs with molecule graphs for precise molecular property predictions.
- Developed TxBKG, a software that automatically extracts KGs from unstructured text using GPT-3.5, actively used by approximately 100 MD students in UIUC. [link] [code]

Research Assistant to Prof. Jiawei Han Sep. 2022 – May 2024 (expected)

- Developed TAGREAL, a framework mining optimal prompts from large corpora for the KG completion task with pre-trained language models (PLMs).
- Developed TRISUM, a method training student model with the rationales from LLMs for abstractive summarization. We introduced a novel scoring mechanism for golden rationale selection and a phased curriculum learning for effective knowledge distilling.
- Introduced GENRES, a multi-aspect framework to evaluate the generative relation extraction capabilities of large language models.
- (In progress) Developing GOLDMINER, an adaptive prompt mining method that aligns prompt generation LMs with knowledge base LMs through direct preference optimization (DPO) for KG completion task.

Teaching Assistant (CS598: Deep Learning for Healthcare) to Prof. Jimeng Sun Jan. 2023 – May 2023

- Provided one-on-one assistance to students on course material and programming assignments. Taught students the insights of popular deep learning (DL) models, and guided them to build their own pipelines of DL for health with PyHealth. Graded students' project reports and presentations.

Teaching Assistant (CS101: Intro Computing) to Prof. Mattox Beckman Aug. 2022 – Dec. 2022

- Led lab sections and office hours, designed over 200 questions for examinations, and taught fundamentals of data structures and algorithms, programming languages like Python and MATLAB to students.

Massachusetts Institute of Technology Cambridge, USA

Research Projects with Prof. Mark Vogelsberger May 2021 – Aug. 2021

- Developed a skin lesion diagnosis software applying an ensemble of convolutional neural network-based models voted by cubic precision.

Waseda University Tokyo, Japan

Research Assistant to Prof. Shinichi Honiden and Prof. Kenji Tei Jun. 2020 – Jul. 2021

- Developed OACAL, an algorithm combining model checking and machine learning techniques to automatically revise software specifications.
- Researched on the joint approach of machine learning and formal verification.

Teaching Assistant (Algorithms and Data Structures) to Prof. Honiden Shinichi May 2020 – Sep 2020

- Led lab sections and office hours teaching classic algorithms and data structures to students.

Industrial Experience

Relativity Chicago, USA

PhD Research Intern May. 2023 – Aug. 2023

- (1) Researched leveraging rationales retrieved from LLM as additional supervision to train student model for text summarization. (2) Introduced a graph pooling-based method for long document summarization.

Ant Group (Alibaba) Hangzhou, China

Software Engineer (AI Engineering) Nov. 2021 – Jul. 2022

- (1) Developed a high-performance machine learning system (MLOps) named AlphaRisk for diverse industries to use (mainly used by Alipay's intelligent security system). (2) Developed Auto-Refit for Deep Learning, a tool embedded in AlphaRisk enabling dynamic update of the DL models with new features.

Skills & Services

- **Programming Languages & Tools:** Python, Java, JavaScript, Fortran, CSS, MATLAB, Typescript, C/C++, PHP, R, Golang, MySQL. PyTorch, PyTorch Geometric, DGL, TensorFlow, Keras, TuriCreate, MySQL, Hadoop, Spring, Git, Linux, Model Checking tools.
- **Paper Review Services:** EMNLP'23, AAAI'24, NAACL'24