Patrick (Pengcheng) Jiang

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

Education

University of Illinois at Urbana-Champaign — Champaign, USA
M.S. in Computer Science (research-based program, advisor: Prof. Jimeng Sun) — Sep. 2022 – May 2024 (expected)

- GPA: 4.0/4.0 (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 — Sep. 2017 – Sep. 2021

- GPA: 3.86/4.0 (Rank: Top 2%); 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, Adam Cross, and Jimeng Sun, "GraphCare: Enhancing Healthcare Predictions with Open-World Personalized Knowledge Graphs", (under review at *NeurIPS'23*), 2023. [pdf]
- **Pengcheng Jiang**, Shivam Agarwal, Bowen Jin, Xuan Wang, Jimeng Sun, and Jiawei Han, "Text-Augmented Open Knowledge Graph Completion via Pre-Trained Language Models", *2023 Annual Meeting of the Association for Computational Linguistics* (*ACL'23*), 2023. [pdf]
- Chaoqi Yang, Zhengbang Wu, **Patrick Jiang**, Zhen Lin, Junyi Gao, Benjamin Danek, and Jimeng Sun, "PyHealth: A Deep Learning Toolkit for Healthcare Predictive Modeling", *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (*KDD'23*), 2023. [link]
- **Pengcheng Jiang**, Cao Xiao, Tianfan Fu, and Jimeng Sun, "Bi-level Contrastive Learning for Knowledge-Enhanced Molecule Representations", (in submission), 2023. [pdf]
- **Pengcheng Jiang**, Cao Xiao, Zifeng Wang, Jimeng Sun, and Jiawei Han "TriSum: Learning Summarization Ability from Large Language Models", (under review at *AAAI'24*), 2023.

Undergraduate Research

- **Pengcheng Jiang**, Kenji Tei, "OACAL: Finding Module-consistent Specifications to Secure Systems from Weakened User Obligations", *IEEE Symposium Series on Computational Intelligence* (*SSCI*), 2021. [pdf]
- **Pengcheng Jiang**, "CNN-based Diagnosis System on Skin Cancer using Ensemble Method Weighted by Cubic Precision", *TechRxiv Preprint*, 2021. [pdf]

Academic Experience

University of Illinois at Urbana-Champaign — Champaign, USA
Research Projects (as a Research Assistant since Aug. 2023) with Prof. Jimeng Sun — Sep. 2022 – Expected May 2024

- Developed PyHealth, a comprehensive deep learning toolkit designed for healthcare. [link].
- Developed GraphCare, an innovative framework that leverages open-world knowledge graphs (KGs) from Large Language Models (LLMs) and established KGs to create patient's personalized KGs for enhanced clinical predictions.
- Developed TextbookKG, a software powered by LLMs to craft KGs from unstructured text. [link].
- Developed Gode, a novel approach that synergizes large-scale biochemical knowledge graphs with molecule graphs for precise molecular property predictions.
- Introduced MedKG, a cutting-edge method that harnesses textbook-derived medical knowledge to enrich the learning experience of medical students.

Research Projects with Prof. Jiawei Han Sep. 2022 – Feb. 2023

- Developed TagReal, a framework leveraging pre-trained language models (PLMs) on open knowledge graph completion tasks. TagReal bridges traditional text mining and contemporary prompt mining.
- Developed TriSum, a method distilling rationales from LLMs for abstraction summarization tasks. We use a topic-distribution-based method (LDA) for the golden rationale selection process.

Teaching Assistant (CS 598: Deep Learning for Healthcare) to Prof. Jimeng Sun Jan. 2023 – May 2023

- Led office hours. Taught students state-of-the-art AI models, and guided them to use PyHealth to build their own pipelines of deep learning for health. Graded students' project reports and presentations.

Teaching Assistant (CS 101: 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, Python, and MATLAB.

Massachusetts Institute of Technology Remote

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.

Professional Experience

Relativity Chicago, USA

PhD Research Intern May. 2023 – Aug. 2023

- Researched on leveraging rationale graphs retrieved from LLM to train a local lightweight Seq2Seq model (knowledge distillation) for text summarization.
- Introduced a graph pooling-based method for controllable long document summarization.

Alibaba - Ant Group Hangzhou, China

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

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

Tencent Shenzhen, China

Research Engineer Intern Sep. 2021 – Oct. 2021

- Developed a multi-label classification and sentiment analysis application using 5 DL algorithms (textCNN, DPCNN, BiLSTM, Bert, RoBERTa) and 3 traditional ML algorithms (LR, SVM, XGBoost).

Skills

- **Programming Languages:** Python, Java, JavaScript, CSS, Fortran, MATLAB, Typescript, C/C++, PHP, R, Golang, MySQL.
- **Tools/Frameworks:** PyTorch, PyTorch Geometric, DGL, TensorFlow, Keras, TuriCreate, MySQL, Hadoop, Spring, Git, Linux, model-checking tools.