Xiaoda Wang

ignite-abd.github.io • ignite-abd • awangxiaoda@gmail.com

Education

Department of Cyber Science and Engineering, Sichuan University(SCU) **Sep 2020 - Jun 2024**

B.Eng in Cybersecurity *Chengdu, China*

- GPA: 3.88/4, Rank: 4/194 (**2.06%**)
- Advisor: Prof. Mingjie Tang (obtained Ph.D. degree at Purdue University) and Prof. Haizhou Wang
- A+ Courses: Calculus, Data Structures and Algorithmics, Computer Architecture, and 38 others

Research Interests

I am a highly motivated and passionate undergraduate eagerly seeking admission to a Ph.D. program. My principal research interests are generally in **Data Mining**, **Graph Neural Networks**, **Large Language Models**, **Machine learning System**, and **AI Security**. My ultimate goal is to drive meaningful advancements in research that have a tangible impact on real-world applications.

Publications

- **Xiaoda Wang**, Zhaoyi Liu, Tengda Guo, Zhiyuan Cheng, Carl Yang and Mingjie Tang, "BadDet: Blockchain Fraud Detection with Dynamic Address-Transaction Graph Convolutional Networks" in International World Wide Web Conferences (**WWW**), 2024. (In Submission)
- **Xiaoda Wang**, Yuan Tang, Tengda Guo, Bo Sang, Jingji Wu, Jian Sha, Ke Zhang, Jiang Qian, Mingjie Tang, "Couler: Unified Machine Learning Workflow Optimization in Cloud" in International Conference on Data Engineering (**ICDE**), 2024. (In Submission)
- **Xiaoda Wang**, Tengda Guo, Chenhui Hu, Mingjie Tang, and Carl Yang, "Enhancing Large Language Models with Knowledge Graphs and Graph Neural Networks" in International Joint Conference on Artificial Intelligence (**IJCAI**), 2024. (In Pre-submission)
- **Xiaoda Wang**, Chenxiang Luo, Tengda Guo, Zhangrui Liu, Jiongyan Zhang, and Haizhou Wang, "BGEK: External Knowledge-enhanced Graph Convolutional Networks for Rumor Detection in Online Social Networks," in International Conference on Artificial Neural Networks (**ICANN**), 2023.
- Yifei Jian, Xinyu Chen, **Xiaoda Wang**, Ying Liu, Xingshu Chen, Xiao Lan, Wenxian Wang, and Haizhou Wang, "A Metadata-aware Detection Model for Fake Restaurant Reviews based on Multimodal Fusion," in Information Processing and Management (**IPM**), 2023. (In Submission)

Research Experience

Blockchain Fraud Detection with Dynamic Address-Transaction GCN **Dec 2022 - Oct 2023**

Advisors: Prof. Carl Yang (Emory University); Prof. Mingjie Tang (SCU) *Research Assistant*

- Converted Bitcoin transactions into an address-transaction graph structure, creating the first large-scale dynamic heterogeneous Bitcoin dataset with 47 time steps and over 850k Bitcoin addresses.
- Developed a clustering algorithm for user entity graph construction and user entity feature extraction.
- Proposed a dynamic GCN utilizing an unsupervised feature generation approach to obtain low-dimensional representations, aiding in effective fraud identification. The dynamic GCN updates the weight matrices of different layers along the temporal dimension.
- Finished paper: a paper in submission to WWW 2024.

Couler: Unified Machine Learning Workflow Optimization in Cloud **Apr 2023 - Oct 2023**

Advisors: Prof. Mingjie Tang (SCU) *Research Assistant*

- Provided a unified programming interface for workflow definition, ensuring independence from the workflow engine and compatibility with various workflow engines.
- Integrated Large Language Models in unified programming code generation using natural language descriptions and automated hyperparameter tuning through Dataset Card and Model Card integration.

- Divided large workflows into smaller ones for auto-parallelism optimization and implement dynamic artifact caching to minimize redundant computations and ensure fault tolerance.
- Finished paper: a paper in submission to ICDE 2024.

Enhancing Large Language Models with Knowledge Graphs and GNN **Apr 2023 - Now**

Advisors: Prof. Carl Yang (Emory University); Prof. Mingjie Tang (SCU) *Research Assistant*

- Developed a plug-and-play prompting approach to elicit a graph-of-thoughts question-answering capability in Large Language Models.
- Introduced a pruning method based on GNN and Contrastive Learning to retrieve the most relevant subgraph from the Knowledge graph.
- Proposed a graph-to-text conversion technique to enhance understanding of graph structures and implemented self-calibration to improve the confidence scores.
- Finished paper: a paper in pre-submission to IJCAI 2024.

BGEK: External Knowledge-enhanced GCN for Rumor Detection **Oct 2022 - Apr 2023**

Advisors: Prof. Haizhou Wang (SCU) *Research Assistant*

- Constructed the first structural Cantonese rumor dataset containing source tweets, retweets and comments in social networks.
- Developed a novel method for extracting external knowledge features based on GCN and obtaining correlation features through a Comparison Network.
- Proposed a novel Cantonese rumor detection model BGEK for Cantonese rumor detection, which integrates the text features, comparison features, and structural features of tweets.
- Finished paper: a paper to ICANN 2023.

A Metadata-aware Detection Model for Fake Restaurant Reviews **Apr 2022 - Jul 2023**

Advisors: Prof. Haizhou Wang (SCU) *Research Assistant*

- Constructed the first publicly available dataset of fake restaurant reviews with extensive metadata;
- Introduced 10 brand-new features and 9 redefined features to complement the feature engineering of existing works and improved the detection performance, which are derived from both review text and metadata;
- Developed a novel metadata-aware model to identify fake restaurant reviews based on multimodal fusion, learning the contextualized representation and emotional tendencies in review text, while effectively integrating abundant metadata;
- Finished paper: a paper in submission to IPM 2023.

Experience

Innovation and Interdisciplinary Research Programme **Jan 2023 – Feb 2023**

Advisors: José Miguel Hernández-Lobato; Grade: A *University of Cambridge*

- Financial Fraud Detection using Graph Neural Networks.
- Interdisciplinary thinking, research methods, scientific literacy and professional knowledge.

Selected Awards

- Outstanding Student Award 2021, 2022, 2023
- Excellent Student Cadre Award 2022
- Second-Class Comprehensive Scholarship 2022, 2023
- Third Prize Winner, China Undergraduate Mathematical Contest in Modeling 2022
- First Prize Winner, Chinese Physics Olympiad (CPHO) 2019

Skills

Programming Languages: Python, C

Tools and Frameworks: LaTeX, PyTorch, TensorFlow, PyG, Keras, Docker, Linux