Yiyue Qian

https://yiyueqian.github.io/

yqian5@nd.edu
+1-216-370-8373

Experience

- **Case Western Reserve University** — Cleveland, USA
  Research Assistant (Full-time) — *Sept. 2019 - present*
  - **Drug Trafficking Detection:** To detect drug traffickers on *social media*, developed a novel framework MetaHG by integrating graph representation learning and meta-learning over the built heterogeneous graph which incorporates multi-modal data and relational structured information among nodes in graph. (93% of drug traffickers in our dataset can be detected.)

    To study drug trafficking among different *darknet markets*, developed an adversarial framework dStyle-GAN that combines writing and photography styles of drug posts, graph embedding learning, and adversarial learning to identify the same drugs posted among different darknet markets. (89.3% of posted drug pairs among multiple darknet markets in our dataset can be identified correctly.)
  - **Malicious Repository Detection:** Built a novel model Meta-AHIN integrating GCN, masking-based self-supervised learning, and task-aware meta-learning to detect COVID-19 themed malicious repositories over the heterogeneous information network using *GitHub* data. (87.3% of malicious repositories in our dataset can be successfully detected.)
  - **α-Satellite Website for Assessing Risk during COVID-19:** Developed a website called α-Satellite with a dynamic map and an analysis board to provide a risk index at multiple levels (e.g., county-level) and information on public perception about the pandemic; Built a model to learn the area embedding based on the heterogeneous graph augmented by cGAN for calculating the risk index. (more than 50k visits https://covid-19.yes-lab.org)
- **Ping An Technology Co., Ltd.** — Shenzhen, China
  Algorithm Engineer (Full-time) — *Jan. 2018 - Aug. 2019*
  - **Model Building:** Built NLP-related models to capture main topics, recognize named entities, analyze relationships among entities, and judge sentiment from unstructured text for Ping An FinTech AI analysis platform.
  - **Risk Event Monitoring:** Developed models on Linux to analyze the topic of global daily financial news such as Reuters News and monitor risk events at different levels (bankruptcy, violation of laws, affairs of executives, etc.). This launched function on the AI analysis platform is estimated to save more than $1 million per year for Ping An Group.
- **China Vanke Co., Ltd.** — Guangzhou, China
  Data Analyst (Full-time) — *Oct. 2017 - Dec. 2017*
  - **Model Building:** Built machine learning models to analyze impact factors of commercial buildings and predict trends in real estate prices in big cities.
  - **Marketing Analysis:** Visualized results in Python for marketing insights and analyzed the trend on the model results.

Education

- **University of Notre Dame** — South Bend, USA
  PhD in Computer Science — *Jan. 2022 - Jun. 2024*
  Research Interest: Graph Representation Learning, Self-Supervised Learning, Machine Learning, Deep Learning, Data Mining
- **Case Western Reserve University** — Cleveland, USA
  PhD in Computer Science — *Sept. 2019 - Dec. 2021*
  Research Interest: Graph Representation Learning, Self-Supervised Learning, Machine Learning, Deep Learning, Data Mining
- **University College London** — London, UK
  Master in Statistics — *Sept. 2016 - Sept. 2017*
  Courses: Machine Learning, Statistical Models and Data Analysis, Applied Bayesian Methods, Statistics Inference, Decision and Risk
  Thesis: Classifying Three Propensities in Orange Telecom Company

Publication

- **Qian, Y.**, Zhang, Y., Ye, Y. and Zhang, C.. Distilling Meta Knowledge on Heterogeneous Graph for Drug Trafficker Detection on Social Media, **NeurIPS**, 2021
- **Qian, Y.**, Zhang, Y., Ye, Y. and Zhang, C.. Adapting Meta Knowledge with Heterogeneous Information Network for COVID-19 Themed Malicious Repository Detection, **IJCAI**, 2021.
- Zhang, Y., **Qian, Y.**, Fan, Y., Ye, Y., Li, X., Xiong, Q. and Shao, F.. dStyle-GAN: Generative Adversarial Network based on Writing and Photography Styles for Drug Identification in Darknet Markets, **ACSAC**, 2020.
- Ye, Y., Fan, Y., Hou, S., Zhang, Y., **Qian, Y.**, Sun, S., Peng, Q., Ju, M., Song, W. and Loparo, K.. Community Mitigation: A Data-driven System for COVID-19 Risk Assessment in a Hierarchical Manner. **CIKM**, 2020.

Skills Summary

- **Languages and Platforms** Python (PyTorch, TensorFlow, Keras), R, C++, JavaScript, SQL, Linux, Colab
- **Soft Skills** Leadership, Team-Spirited, Problem Solving, Motivated, Critical Thinking, Creative Thinking
- **Languages** Mandarin, English