ZICUN CONG

Burnaby, B.C., Canada, V5A 0A8

📞 +1 (604) 7825610   ✉ congzicun@gmail.com   ☁ minzc.github.io   ⭘ github.com/Minzc

Education

Simon Fraser University   Jan. 2017 – Sep. 2022

Ph.D. in Computing Science, GPA: 4.0/4.33, Supervisor: Prof. Jian Pei

Thesis: "Towards Trustworthy Data Analytics: Algorithmic Tools for Interpretability and Fairness"

Industry Experience

Zscaler Inc.   Nov. 2022 – Present

Staff Data Scientist   *Burnaby, Canada*

- Led a team of data scientists, engineers, and security researchers to develop robust data pipelines and innovative graph-based models for cyber threat detection on the Google Cloud Platform (GCP)
- Spearheaded the innovation and development of advanced interpretation tools for ML threat detection models, resulting in a 100% reduction in security researcher verification time
- Developed heterogeneous graph neural networks using Python to detect malicious campaigns, achieving a 35% improvement in accuracy over the legacy system
- Optimized the legacy data pipeline, resulting in a 25% reduction in computation time

Fortinet Inc.   Sep. 2016 – Nov. 2022

Staff Software Engineer   *Burnaby, Canada*

- Developed advanced machine learning and data mining models using Python libraries such as Pandas, NumPy, Scikit-Learn, and PyTorch to detect malicious traffic and protect millions of endpoints, increasing accuracy by up to 20%
- Designed and implemented robust data pipelines using PySpark and Hadoop to process billions of network traffic and malware behavior logs, and extract features stored in HBase for further analysis
- Delivered a NLP service in Python to automatically extract open-source threat intelligence from unstructured text, enabling faster and more efficient threat analysis

LinkedIn.   May. 2021 – Dec. 2021

Machine Learning Engineer Intern   *Remote*

- Implemented several bias mitigation techniques to eliminate unfairness in Graph Neural Networks, improving the fairness and reliability of the models
- Conducted research into the problem of unfairness in graph neural networks, proposing an innovative sampling framework that uses reinforcement learning to improve fairness and retain accuracy

Selected Publications

- Xuan Luo, Jian Pei, **Zicun Cong**, Cheng Xu. "On Shapley Value in Data Assemblage Under Independent Utility." Proc. VLDB Endow. 15, 11 (2022), 2761–2773.

- **Zicun Cong**, Xuan Luo, Jian Pei, Feida Zhu, Yong Zhang. "Data Pricing in Machine Learning Pipelines." Knowledge and Information Systems (KAIS), 2022.

- Jian Pei, Feida Zhu, **Zicun Cong**, Xuan Luo, Huiwen Liu, Xin Mu. "Data Pricing and Data Asset Governance in the AI Era." In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD '21). Association for Computing Machinery, New York, NY, USA, 4058–4059.

- **Zicun Cong**, Lingyang Chu, Yu Yang, Jian Pei. "Comprehensible counterfactual explanation on Kolmogorov-Smirnov test." Proc. VLDB Endow. 14, 9 (May 2021), 1583–1596.

- **Zicun Cong**, Lingyang Chu, Lanjun Wang, Xia Hu, and Jian Pei. "Exact and Consistent Interpretation of Piecewise Linear Models Hidden behind APIs: A Closed Form Solution." In 2020 IEEE 36th International Conference on Data Engineering (ICDE), pp. 613-624. IEEE, 2020.

Technical Skills

Artificial Intelligence: Deep Learning, Machine Learning, Data Mining
Languages and Frameworks: Python, Java, C++, SQL, PyTorch, TensorFlow, Sklearn
Cloud and Distributed Computing: Hadoop, Spark, HBase, Phoenix db, Redis, Mysql, Zookeeper, Apache Storm