

Zhiyi Chen

Tel: +1 (470)659-3346 Email: zchen798@gatech.edu
Klaus Advanced Computer Building, room 3337, Atlanta, GA, US

Education

Georgia Institute of Technology

Ph.D. Candidate in Computer Science

GPA: 4.0/4.0

Atlanta, GA

Jan 2021 - present

Tsinghua University

Bachelor of Science in Computer Science and Technology

GPA: 3.73/4.0

Beijing, China

Aug 2016 - Jul 2020

Research Experience

Internet Measurement: Internet Autonomous Systems Tagging Framework

Georgia Institute of Technology

Research Assistant; Advisor: Alberto Dainotti

Dec 2022 - present

- Designing a framework/toolkit for providing user-customized tags that capture key properties of Autonomous Systems.

Internet Measurement: Improving the Inference of Sibling Autonomous Systems

Georgia Institute of Technology

Research Assistant; Advisor: Alberto Dainotti

Aug 2021 - Dec 2022

- Identified two root causes that led to systematic inaccuracies in the state-of-the-art AS-to-organization dataset.
- Built an improved dataset on sibling relations, correcting the mappings of 12.5% of organizations with sibling ASes.

Machine Learning: Automatic Tabular Data Preprocessing Pipeline

Georgia Institute of Technology

Research Assistant; Advisor: Xu Chu

Jan 2021 - Aug 2021

- Proposed a method to automatically and efficiently search for a data pre-processing pipeline for a given tabular dataset while maximizing the downstream machine learning model at the same time.
- Achieved the best test accuracy on 15 real-world datasets and improved the model's test accuracy by up to 6.6%.

Machine Learning: Continual Learning Algorithm Design

Peking University

Research Intern; Advisor: Tong Lin

Apr 2019 - Feb 2021

- Proposed a novel gradient descent training framework for mitigating catastrophic forgetting in continual learning.
- Improved the average test accuracy by up to 7.9% and reduced forgetting by up to 15.4% on four datasets.

Work Experience

Coalition, Inc.

Remote

Research Intern; Advisor: Daniel Woods, Morgan Hervé-Mignucci

Jun 2023 - Aug 2023

- Leveraged public Internet measurement platforms and GIS tools to infer and visualize geographic physical locations of datacenters that host policyholder domains.
- Explored the aggregation of cyber risks on physical infrastructure, especially critical points such as datacenters.

Publications / Presentations

- Improving the Inference of Sibling Autonomous Systems**, published at the Passive and Active Measurement Conference (PAM) 2023.
 - Our dataset won *PAM 2023 Best Community Artifact Award*.
 - A lightning talk presented at NANOG (North American Network Operators' Group) 87 in Atlanta.
 - A blog posted at APNIC (Asia Pacific Network Information Centre) Blog website.
- DiffPrep: Differentiable Data Preprocessing Pipeline Search for Learning over Tabular Data**, published at The ACM Special Interest Group on Management of Data (SIGMOD) 2023.
- Principal Gradient Direction and Confidence Reservoir Sampling for Continual Learning**, published at the International Conference on Artificial Neural Networks (ICANN) 2021.

Skills

- Programming language: Python (PyTorch), C++, SQL, P4
- Technical skills: machine learning algorithms, deep learning models, Internet databases, data analytics