Zhi Yang

SUSTech,1088 Xueyuan Avenue, Nanshan District, Shenzhen 518055, P.R. China
13828741098|12332454@mail.sustech.edu.cn

EDUCATION

Master's degree in Electronic Science and Technology Expected June 2026
Southern University of Science and Technology

- Master's Thesis: The Combination of Differential Privacy and Fairness in ML

Bachelor's degree in Software Engineering June 2023
Yunnan University

- GPA: 3.8/4.0

PUBLICATIONS

Zhi Yang, Changwu Huang, Xin Yao. **Towards Private and Fair Machine Learning: Group-Specific Differentially Private Stochastic Gradient Descent with Threshold Optimization**. The International Conference on Neural Information Processing, 2024. (Recently Accepted)

- Existing studies primarily address fairness issues in differentially private models in isolation. To bridge this gap, we propose a group-specific DP stochastic gradient descent (DP-SGD) training mechanism with classification threshold optimization, which simultaneously tackles both accuracy and outcome fairness challenges in differentially private models.

Zhi Yang, Ziming Wang, Changwu Huang, Xin Yao. **An Explainable Feature Selection Approach for Fair Machine Learning**. International Conference on Artificial Neural Networks. Cham: Springer Nature Switzerland, 2023: 75-86.

- In this work, we propose an explainable feature selection (ExFS) method to improve the fairness of ML by recursively eliminating features that contribute to unfairness based on the feature attribution explanations of the model's predictions.

AWARDS AND HONORS

Excellent teaching assistant	September 2024
Top research assistance reward	September 2024
National Scholarship	December 2022
The 8th Internet plus Innovation and Entrepreneurship Competition, National Silver Award	November 2022
University-level first-class scholarship	December 2021
University-level first-class scholarship	December 2020

PROJECT EXPERIENCE

Analysis of the Concept of Transparency in Intelligent AI Systems September 2023
SUSTech Trusted Intelligent Systems Innovation Laboratory

- Summarizing and analyzing existing concepts of transparency by drawing insights from articles on AI ethics, societal implications of AI, legal regulations.

Other projects:

- *Fuzzy Facial Processing and Recognition System:* Conducted an extensive review of the latest advancements in low-resolution facial recognition, selecting suitable approaches for data preprocessing and model training.
- *Facial Recognition System Based on Tracking Algorithms:* Integrated the Hongsoft facial recognition interface and developed the system using the Springboot, Mybatis-Plus, and Angular frameworks.
- *Huan Yan software:* Leveraged the open-source GFPGAN algorithm to restore old and blurred photos, developed on the WeChat Mini Program platform using WXML, WXS, and JavaScript.