Jiao Sun

jiaosun.thu@gmail.com sunjiao123sun.github.io

Research Interest: Human-Computer Interaction, Machine Learning, Data Visualization

EDUCATION

- **M.Sc., Computer Science and Technology, IIIS, Tsinghua University, Beijing, China** 09/2016 - 07/2019
  - Overall GPA: 3.7/4.0, Rank 1/6 (all master students in the department); Advisor: Wei Xu
- **B.E., Computer Science Department, Shandong University, China** 08/2012 - 06/2016
  - Overall GPA: 4.0/4.0; Rank 1/43 (graduated with the highest honor)

PUBLICATION

- **Jiao Sun,** Yin Li, Lei Shi, Charley Chen, Ling Huang, and Wei Xu, "FDHelper: Assist Unsupervised Fraud Detection Experts with Interactive Feature Selection and Evaluation" (Under Review at **SIGCHI 2020**) [initial review 3.5/5]
- Yikun Ban, **Jiao Sun,** Xin Liu, Ling Huang, Yitao Duan, and Wei Xu, "FraudTrap: Catching smart Fraudsters via Object Similarity Graph Analysis" (Arxiv Preprint)
- **Jiao Sun,** Qixin Zhu, Zhifei Liu, Xin Liu, Yueming Wang, Jihae Lee, Lei Shi, Ling Huang and Wei Xu, "FraudVis: Understanding Unsupervised Fraud Detection Algorithms" (**PacificVis 2018**)
- Charley Chen, Guosai Wang, **Jiao Sun** and Wei Xu, "Detecting Data Center Cooling Problems Using a Data-driven Approach", the 9th Asia-Pacific Workshop on Systems (**APSys 2018**)
- Yang Zhang, Tingjian Zhang, Yongzheng Jia, **Jiao Sun** and Wei Xu, "DataLab: Introducing Software Engineering Thinking into Data Science Education at Scale" (**ICSE 2017**)

AWARDS

- Tsinghua-RWTH Aachen University Research Fellowship 10/2018
- Baidu-Tsinghua Future Star Scholarship (top 3%) 10/2017
- Outstanding Graduate of Shandong Province (top 1%) 05/2016
- Outstanding Graduate of Shandong University (top 3%) 05/2016
- First-Class Scholarship of Shandong University (consecutive three times, top 2%) 2013~2016
- Merit Student of Shandong University (consecutive three times, top 2%) 2013~2016
- National Motivational Scholarship (consecutive three times, top 2%) 2013~2016
- Excellent Student Cadre of Shandong University (top 3%) 06/2016
- First Prize, Mathematical Contest of Modeling in Shandong University (top 1%) 08/2015
- Bronze Award, National Mobile Innovation Programming Contest 05/2014
- Best debater of Shandong University 03/2013

SKILLS

- **Proficient** in Python, Matlab, Java, Django, Linux, Shell, SQL, Neo4j
- **Familiar** with Tensorflow, PyTorch, Keras, Javascript, R, Html, D3.js, Spark, Hadoop, GPU Programming, R

RESEACH & PROJECTS

- **FDHelper: Assist Unsupervised Fraud Detection Experts with Interactive Feature Selection and Evaluation [1]**
  KEYWORDS: Human Computer Interaction; Interactive ML assistant tool; Fraud Detection; Visualization;
  Registered three national-level patents in China; Co-authored with Prof. Wei Xu and Andrew Chi-Chih Yao;
  proposed an interactive machine learning assistant tool with the visual guidance for fraud detection;
  designed a scalable, multi-layer map to interpret the quality of fraud detection algorithms;
  proposed a novel metric to consolidate the sub-space grouping characteristics of fraud users;
  conducted user studies and semi-structured interviews to demonstrate the effectiveness.

- **FraudTrap: Catching Smart Fraudsters via Object Similarity Graph Analysis [2]**
  KEYWORDS: Anomaly Detection; Graph Algorithm; Machine Learning; Data Mining
  Built Object Similarity Graph using a novel similarity metric to catch loosely synchronized behavior of fraud groups;
  designed a scalable and provably converging algorithm implemented in Apache Spark;
  the algorithm works for both the unsupervised and semi-supervised modes.

- **FraudVis: Understanding Unsupervised Fraud Detection Algorithms [3]**
  KEYWORDS: Anomaly Detection; Data Visualization; Human Computer Interaction
  Won the **Best Poster Award** in the 8th Cross-Strait Tsinghua Postgraduate Academic Forum;
  proposed a comprehensive data visualization system for understanding the fraud detection result;
  analyzed the fraud characteristics from inter-group, intra-group, individual and temporal perspectives;
  customized different interactions for different target users.

- **DataLab: Introducing Software Engineering Thinking into Data Science Education at Scale [5]**
  KEYWORDS: Software Engineering; Education; Large-scale Computation
  The system is still in use for the Big Data System course in Tsinghua University
  Designed and implemented a web-based system for the machine learning education;
  provided a platform for machine learning researchers to manage their code and data together;
  provided a leaderboard for students to better evaluate their result with others;
  helped teachers to better understand and analyze both the course difficulty level and students' study status;
  provided a hands-on online lab environment to train students with a focus on the data science contents.

- **Research Intern at Informatik 5 Lab, RWTH Aachen University, Germany**
  KEYWORDS: Anomaly Detection; Data Generation; Machine Learning; Data Mining
  Got Tsinghua-RWTH research fellowship to work on the project;
  used machine learning based models on detecting anomalies from the entry data in Bosch company;
  generated synthetic data to improve the generality of the model.

WORK EXPERIENCE

- **Data Science Intern at Tencent, Beijing** 05/2019-07/2019
  KEYWORDS: Video Quality Test; Software Engineering; Data Mining
  Got selected to 2019 Tencent Elites Internship Plan;
  Analyzed Tencent *WeSee* in comparison with its competitor product *TikTok;*
  surveyed about the current video compression methodologies;
  implemented interfaces for crowdsourcing platform and did user studies about the design of the platform;
  proposed and implemented new interactions for users to manually label the subjective comment for two videos.