Xiao Hui Tai

✉ xtai@berkeley.edu 📞 412-557-2835 ⌨ xhtai 🌐 xhtai.github.io
Updated: May 2, 2020

Academic Appointment

University of California, Berkeley, School of Information
Postdoctoral Researcher   Jan 2020 - Present

Education

Carnegie Mellon University
Ph.D. in Statistics   Aug 2015 - May 2019
Thesis: *Forensic Data Matching Problems*

Master of Statistical Practice   Received May 2012

Singapore Management University
B.Sc. in Economics, *summa cum laude*   Received May 2011

Publications

Xiao Hui Tai, Kyle Soska and Nicolas Christin. Adversarial Matching of Dark Net Market Vendor Accounts. *25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2019.

Xiao Hui Tai. Record Linkage and Matching Problems in Forensics. *IEEE 18th International Conference on Data Mining Workshops (ICDMW)*, 2018.

Xiao Hui Tai and William F. Eddy. A Fully Automatic Method for Comparing Cartridge Case Images. *Journal of Forensic Sciences*, 2018.

Pre-prints

Xiao Hui Tai and Kayla Frisoli. Benchmarking Minimax Linkage in Hierarchical Clustering. *Springer Series "Studies in Classification, Data Analysis, and Knowledge Organization"* (pending minor revisions), 2020.

Xiao Hui Tai and William F. Eddy. Automatically matching topographical measurements of cartridge cases using a record linkage framework.

Book Chapters and Magazine Articles

Xiao Hui Tai, Open Forensic Science in R, Firearms: casings. 2019.

Alicia Carriquiry, Heike Hofmann, **Xiao Hui Tai** and Susan VanderPlas, Machine learning in forensic applications. Significance, 2019.

Professional Experience

Quantitative Modeler, JPMorgan Chase & Co.   June 2019 - Jan 2020
Develop and maintain models to forecast movements between balance tiers in checking accounts, for use in stress testing.

Statistician, Ministry of Trade and Industry, Singapore   Oct 2012 - August 2015
Maintained a database of demographic information and collaborated with economists and social scientists on policy-related statistical analysis.

Awards and Honors

2019   Honorable Mention for Cluster Benchmark Challenge at the 16th Conference of the International Federation of Classification Societies
2019   Center for Statistics and Applications in Forensic Evidence Poster Award
2017   American Statistical Association Travel Award
2011   Monetary Authority of Singapore Academic Excellence Award
2011   School Salutatorian Award

Teaching and Advising

Teaching assistant: Introduction to Statistical Inference (Spring 2012), Summer Undergraduate Research Experience (Summer 2016), Modern Regression (Fall 2018), Undergraduate Capstone (Fall 2018)

Presentations

Benchmarking Minimax Linkage in Hierarchical Clustering (Invited talk) 16th Conference of the International Federation of Classification Societies (IFCS), Thessaloniki, Greece. 08/2019.

Adversarial Matching of Dark Net Market Vendor Accounts 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD), Anchorage, Alaska. 08/2019.

Record Linkage and Matching Problems in Forensics IEEE 18th International Conference on Data Mining Workshops (ICDMW), Singapore. 11/2018.

Approaches to Matching Darknet Market Seller Accounts. Joint Statistical Meetings, Vancouver, BC, Canada. 08/2018.

Automatic Comparison of Cartridge Cases. (Invited talk) American Bar Association Ninth Annual Prescription for Criminal Justice Forensics Program, New York, NY. 05/2018.

Comparing Cartridge Breechface Marks: 2D vs 3D. (Invited talk) Center for Statistics and Applications in Forensic Evidence (CSAFE) Center Wide Webinar. 01/2018.

Automatic Comparison of Cartridge Breechface Marks. (Invited talk) Firearms Examiner Workshop, Ames, IA. 12/2017.

Matching Seller Accounts on Online Anonymous Marketplaces. (Invited talk) Cyber Analytics Workshop at CyLab Partners Conference 2017, Pittsburgh, PA. 09/2017.

Matching Cartridge Cases: Similarity Metrics and Weight of Evidence. (Invited talk) Workshop on Role of Data, Databases and Expert Knowledge in Forensic Inference, Pittsburgh, PA. 09/2017.

[Matching Seller Accounts on Online Anonymous Marketplaces](#). Women in Statistics and Data Science Conference 2017, La Jolla, CA. 10/2017.

Matching Seller Accounts on Online Anonymous Marketplaces. (Invited talk) Memorial Invited Special Session: Remembering Stephen E. Fienberg at International Conference on Forensic Inference and Statistics, Minneapolis, MN. 09/2017.

Developing Methods for Comparison of Cartridge Breechface Images. Joint Statistical Meetings, Baltimore, MD. 08/2017.

Automatic Cartridge Case Comparison and Evaluation. Association of Firearms and Toolmarks Examiners Conference, Denver, CO. 05/2017.

[Developing Methods for Comparison of Cartridge Breechface Images](#). (Invited talk) Forensics@NIST, Gaithersburg, MD. 11/2016.

Machine Learning Lunch: The Halloween Forensic Spooktacular. (Invited talk) Joint with Neil Spencer at the CMU Machine Learning Lunch, Pittsburgh, PA. 10/2016.

[Assessing Contagious Spread of Competing Hashtags on Twitter](#). Poster at the Public Exhibition of LARC Posters and Demos, Singapore. 10/2012.

Service and Volunteer Experience

SAT Math Instructor at People's Test Preparation Service, Berkeley (2020). Taught high school students SAT math.

Consultant at Ministry of Trade and Industry, Singapore (2019). Presented research and consulted on machine learning projects.

PhD Mentor at Carnegie Mellon Department of Statistics (2018-2019). Mentored a younger PhD student.

Volunteer for Women in Statistics Conference at Carnegie Mellon (2018-2019)

Reviewer for Journal of Forensic Sciences (2018)

Volunteer for Statistics Without Borders (2018). Provided technical and user documentation support for R scripts developed for Whole of Syria Health Cluster to provide emergency and trauma care, and basic healthcare where it has collapsed.

Applying to Graduate School (2017). Spoke at a panel with current graduate students for Women in Statistics.

Skills

Programming languages: R, SQL, SAS (proficient), Matlab (intermediate), Python, Stata (some experience)
Operating systems: OS X, Linux