

Xingchi (Anthony) Li anthony.li@stat.tamu.edu [in](#) doccstat [📄](#) doccstat [🔗](#) xingchi.li

EDUCATION

Texas A&M University - Department of Statistics **College Station, Texas, USA**
Doctor of Philosophy in Statistics **Expected May/August 2025**
Committee: Xianyang Zhang, Abhishek Chakraborty, Lan Zhou, Yong Chen

Georgia Institute of Technology - College of Computing **Atlanta, Georgia, USA**
Master of Science in Computer Science **August 2020**

Texas A&M University - Department of Mathematics **College Station, Texas, USA**
Master of Science in Mathematics **May 2019**
Committee: Peter Howard, Andrea Bonito, Jeffrey Hart

Beihang University - School of Mathematics and System Science **Beijing, China**
Bachelor of Science: Hua Loo-Keng Mathematics Class (Honored Program of Mathematics) **July 2018**

EXPERIENCE

A/B Testing Experiment Data Processing Pipeline (Lavaflow) Data Cube SQL Test **August 2023**
Google LLC, Core Data *Host: Xin Bai, Anthony Lai, (Raymond Ho, Caden Lu)*

- Design and implement a new Bazel macro to make the Lavaflow data cube testing results more user-friendly.
- Write C++ code generation binary to dynamically process different data formats using Google SQL.
- Allow users to write ad-hoc test SQL queries to query the data cube golden minion (RecordIO) and check the query results instead of reading large diffs in the minion files.
- Experiment on the relationship between data sampling rates and confidence interval calculations.

RulesKeeper: Domain Knowledge Based Rule's Manager **August 2022**
Google LLC, Core *Host: Grishma Chole, Korn Potisuntorn, (Chinmaya Sahoo)*

- Reliably identify the true failures, triage and resolve them within a time threshold for jobs on Scheduling Platform.
- Set up the programmatic rules based on the domain knowledge and the historical trends to provide intelligent monitoring and reduce notification redundancy by 70% – 90%.
- Achieved flexibility for client to choose / control how the service should behave by allowing them to set alerting triggers at more granular level (based on Rolling Time window, Failure densities).
- Achieved increased client reliability by allowing them to view how the alert may look like for a specific set of rules in advance (Adaptive thresholds, Recommendation based on historical data, Seasonality and/or trends checks).

Predictive text embedding through large-scale heterogeneous text networks xingchi.li/PTEexperiments

- Develop a heterogeneous graph integrating word, document, and label relationships, leveraging both labeled and unlabeled data for improved text embedding.

PUBLICATIONS

Double Blind Conference Paper Submitted **ATSTATS 2025**

Segmenting Watermarked Texts From Language Models **September 2024**
Xingchi Li, Guanxun Li, Xianyang Zhang *NeurIPS 2024*

- Conduct theoretical analysis of Type I, Type II errors for randomization tests to detect the presence of watermarks.
- Propose a method to segment text into watermarked and non-watermarked text with theoretical guarantees.

fastcpd: Fast Change Point Detection in R

March 2024

*Xingchi Li, Xianyang Zhang**(Submitted) Journal of Statistical Software*

- Implement fast change point detection algorithm 500x faster than the vanilla Pruned Exact Linear Time algorithm.
- Provide consistent estimator for the variance estimation with change points.

EVIboost for the Estimation of Extreme Value Index under Heterogeneous Extremes

October 2022

*Jiaxi Wang, Yanxi Hou, Xingchi Li, Tiandong Wang**Journal of Data Science*



- Proposed a gradient boosting algorithm to estimate a functional extreme value index with heterogeneous extremes.
- Data-driven procedure that captures complex and dynamic structures in tail distributions.

Impact of Dietary Supplementation of L-Citrulline to Meat Goats During Gestation on Reproductive Performance*(Major Revision) Journal of Animal Science and Biotechnology**M Newton, A Lopez, C Stenhouse, K Hissen, E Connolly, Xingchi Li, L Zhou, G Wu, W Foxworth, F Bazer***CI-Bot: A Hybrid Chatbot Enhanced by Crowdsourcing**

November 2017

*Xulei Liang, Rong Ding, Mengxiang Lin, Lei Li, Xingchi Li, Song Lu**Web and Big Data*

ACADEMIC WORKS

Reviewer for Journal of the American Statistical Association  0009-0006-2493-0853Reviewer for Journal of Machine Learning Research  0009-0006-2493-0853

TA	Fall 2024	STAT 610	Distribution Theory	Daren Cline	Texas A&M Univ
TA	Spring 2024	STAT 654	Stat Computing with R & Python	Sharmistha Guha	Texas A&M Univ
TA	Fall 2023	STAT 335	Principles of Data Science	Trevor Harris	Texas A&M Univ
TA	Spring 2023	ECEN 360	Computational Data Science	Kevin Nowka	Texas A&M Univ
TA	Fall 2022	STAT 335	Principles of Data Science	Trevor Harris	Texas A&M Univ
RA	Spring 2022			Xianyang Zhang	Texas A&M Univ
RA	Fall 2021			Xianyang Zhang	Texas A&M Univ
RA	Summer 2021			Jianhua Huang	Texas A&M Univ
TA	Summer 2021	STAT 651	Statistics in Research I	Alan Dabney	Texas A&M Univ
TA	Spring 2021	STAT 421	Machine Learning	Shahina Rahman	Texas A&M Univ
TA	Fall 2020	STAT 604	Statistical Computations	Faron Kincheloe	Texas A&M Univ
TA	Spring 2020	CS 7641	Machine Learning	Charles Isbell	Georgia Inst of Tech

TECHNICAL SKILLS

R 30k+ LOC **fastcpd** available on CRAN, **abseil** available on CRAN**Python** 20k+ LOC PyTorch**C++** 10k+ LOC FlumeC++**Java** 10k+ LOC Promise Graph, Spanner GoRM, gRPC**JavaScript** 3k+ LOC D3**Bazel** 1k+ LOC

Protocol Buffers, SQL, MATLAB, Hadoop, Spark, Apache Pig, Hive