

Arkajyoti (AJ) Bhattacharjee

✉ bhattacharjee.57@osu.edu 🌐 ArkaB-DS.github.io in arkajyoti-aj

Education

The Ohio State University	Columbus, OH, USA
Ph.D. in Statistics, GPA: 3.91/4.00	Aug. 2022–Aug. 2027 (Expected)
Advisor: Dr. Arnab Auddy	
M.S. in Statistics, Department of Statistics	Aug. 2022–Dec. 2024
Graduate Minor in Computer Science, Department of CSE	Aug. 2024–May 2025
Indian Institute of Technology Kanpur	Kanpur, India
M.S. in Statistics, Department of Mathematics & Statistics	Sep. 2020–Jun. 2022
Presidency University	Kolkata, India
B.S. (Honors) in Statistics, Department of Statistics	Aug. 2016–Jun. 2019

Research Interests

Statistically principled methods for heterogeneous data, with emphasis on differential privacy, nonparametric inference, modal inference and learning, two-sample testing, and memorization in generative models.

Publications

1. Nash, J. C., & **Bhattacharjee, A.** (2023). A Comparison of R Tools for Nonlinear Least Squares Modeling. *The R Journal*, 15(4). DOI: [10.32614/RJ-2023-091](https://doi.org/10.32614/RJ-2023-091).

Manuscripts

1. **Bhattacharjee, A.**, & Auddy, A. (2026). Differentially Private Nonparametric Modal Learning with Applications to Regression and Clustering. *Submitted to the Journal of Machine Learning Research (JMLR)*. [arXiv:2607.29675](https://arxiv.org/abs/2607.29675)
2. **Bhattacharjee, A.**, & Auddy, A. (2026). Matched-Mode Testing: Two-Sample Inference for the Locations of Density Modes. *Submitted to the International Conference on Artificial Intelligence and Statistics (AISTATS 2027)*.
3. **Bhattacharjee, A.**, Chakraborty, A., & Auddy, A. Calibrated Nearest-Neighbor Memorization: Exact Inference and Detectability in Generative Models. *In preparation*.

Honors and Awards

- **Ransom and Marian Whitney Award for Research**, Department of Statistics, The Ohio State University. May 2027
- **Scientific Communication Video Contest**, ASA Columbus Chapter. Oct. 2026
- **1 of 6 Finalists**, ASA Nonparametric Statistics Section Student Paper Competition; invited to present in a JSM 2026 topic-contributed session. Aug. 2026
- **People's Vote Poster Presentation Award**, 2026 ASA Midwest Regional Conference in Statistics and Data Science; selected from 34 poster presentations and awarded a \$100 prize; ASA Midwest Regional Conference Travel Award. May 2026

- **1st Place Poster Award**, 2026 Annual Joint Biostatistics Symposium; 1st of 16 presenters from The Ohio State University, Cleveland Clinic, and Case Western Reserve University. Apr. 2026
- **Academic Excellence Awards (2020, 2021)**, Senate Scholarships and Prizes Committee, Indian Institute of Technology Kanpur; awarded to the top 10% of each department based on prior-year CPI. 2020, 2021
- **SVMCM Scholarship**, Directorate of Public Instruction, Government of West Bengal. Nov. 2019

Research Presentations

1. *Did the Cell States Move?* ASA Columbus Chapter Scientific Communication Video Contest. youtu.be/q0ecam_hOME Sep. 2026
2. *Privacy-Aware Nonparametric Modal Learning with Applications to Regression and Clustering*. JSM 2026 Topic-Contributed Session; ASA Nonparametric Statistics Section Student Paper Competition Finalist Presentation. Aug. 2026
3. *Privacy-Aware Nonparametric Modal Learning*. Poster, 2026 ASA Midwest Regional Conference in Statistics and Data Science. May 2026
4. *Privacy-preserving recovery of heterogeneous treatment-response patterns and cancer subtypes*. Poster, 2026 Annual Joint Biostatistics Symposium. Apr. 2026
5. *Beyond Averages: Typical Outcomes and Natural Groupings Under Differential Privacy*. Selected research talk, Edward F. Hayes Advanced Research Forum, The Ohio State University. Mar. 2026

Research Experience

The Ohio State University, Department of Statistics **Columbus, OH, USA**
Graduate Research Associate *May–Aug. 2025*

- Designed **DP-GRAMS**, a differentially private mode-seeking algorithm for estimating density modes under formal privacy guarantees.
- Analyzed convergence and statistical error for private nonparametric mode recovery.
- Implemented privacy-aware simulation and real-data evaluation pipelines for modal learning, modal regression, and clustering.

The Ohio State University, Department of Statistics **Columbus, OH, USA**
Graduate Student Consultant *May–Jul. 2023*

- Consulted on interdisciplinary research projects across computer science, speech and hearing science, geography, health and rehabilitation sciences, and anthropology; cleaned data, engineered features, selected models, and translated statistical findings for applied research clients.

The Ohio State University, Department of Statistics **Columbus, OH, USA**
Course Project, STAT 8820: Statistical Inference for Dynamical Systems *Aug.–Oct. 2025*

Differentially Private Levenberg–Marquardt. Developed a privacy-preserving nonlinear least-squares method by adapting Levenberg–Marquardt updates for individual-level differential privacy and evaluated it on nonlinear dynamical-system benchmarks.

Teaching Experience

The Ohio State University, Department of Statistics **Columbus, OH, USA**
Graduate Student Instructor *Aug.–Dec. 2024*

Instructor of record for STAT 4202, Introduction to Mathematical Statistics II, an upper-level undergraduate mathematical statistics course; delivered lectures, wrote homework and exams, and supported approximately 100 students.

The Ohio State University, Department of Statistics

Columbus, OH, USA

Graduate Teaching Associate

Aug. 2022–Present

- Supported upper-level undergraduate and graduate courses in mathematical statistics, statistical inference, and statistical learning: STAT 4202, 4301, 4620, 6201, and 7301. Responsibilities included recitations, R labs, tutoring, grading, and student support.
- Supported large-enrollment undergraduate service courses in business statistics, life sciences statistics, and probability and statistics for engineers: STAT 1430, 2480, and 3470.

Industry and Open-Source Experience

JPMorgan Chase & Co., CCB Risk Modeling

Wilmington, DE, USA

Summer Associate, Data Analytics and AI Research Program

Jun.–Aug. 2026

- Developed **PRIMO**, an account-level machine-learning framework for predicting highly zero-inflated, right-skewed revolving balances using nearly 10M observations and 2K+ engineered features, supporting downstream interest-income estimation and credit-line decisioning.
- Built an Optuna-based Bayesian feature-selection pipeline with Tweedie XGBoost, gain and SHAP importance, and correlation filtering, reducing the feature space from 2K+ to 100 features.
- Developed nested XGBoost models spanning a squared-error baseline and distribution-aware Tweedie and hurdle-Gamma objectives; separately tuned each family with 30-trial Bayesian hyperparameter optimization and implemented corresponding PyTorch feature-token Transformers.
- Conducted multi-vintage out-of-time validation across overall and segment-level metrics; the best model reduced interest-income estimation RMSE by 1.4% in Year 1 and 1.9% in Year 2 versus the existing benchmark.

Google Summer of Code 2021, The R Project for Statistical Computing

Remote

Open Source Developer

May–Aug. 2021

- **Project:** [Improvements to `nls\(\)`](#); mentors: Dr. John C. Nash and Dr. Heather Turner.
- Developed `nlsCompare`, an R package for comparing nonlinear least-squares fits through gradient, convergence, and residual-sum-of-squares diagnostics.
- Wrote tests for `nlsj`, an interim package for improving functionality in base R's `nls()` workflow.

Accenture Solutions Pvt. Ltd.

Gurugram, India

Data Science Analyst Intern

May–Jul. 2021

- **Project:** Smart Decision Making Using AutoML
- Built a Python and Dash AutoML GUI for preprocessing, model selection, hyperparameter tuning, diagnostics, and prediction.
- Supported classification, regression, clustering, and anomaly-detection workflows for use cases such as churn prediction, segmentation, and fraud detection.
- Received a pre-placement offer for full-time employment based on internship performance.

Leadership and Professional Service

(Bio)Statistics Graduate Student Association,

Columbus, OH, USA

The Ohio State University

President

Aug. 2026–Present

- Elected president representing statistics and biostatistics graduate students; lead qualifying-exam information sessions, faculty and staff award voting, and student-focused departmental initiatives.

Science Day, Department of Statistics, The Ohio State University

Columbus, OH, USA

Judge, Statistical Analysis Award

Apr. 2024