Chuting Xu

(607) 339-1783 | cx252@cornell.edu | https://chutingx.github.io

EDUCATION

CORNELL UNIVERSITY Ithaca, NY, United States
BOWERS COLLEGE OF COMPUTING AND INFORMATION SCIENCE
Master of Professional Studies in Applied Statistics GPA 3.703 Conferred Dec 2024
Coursework: Probability Models & Inference; Statistical Computing; Linear Models (matrix methods); Applied Time Series; Data Mining & Machine Learning; Big Data Management; Applied SAS Computing.

UNIVERSITY OF CALIFORNIA, LOS ANGELES Los Angeles, United States
Bachelor of Science in Mathematics GPA 3.621 Conferred Sep 2022
Coursework: Real Analysis I–II; Linear Algebra; Optimization; Probability & Mathematical Statistics I–II; Complex Analysis; Abstract Algebra.
Recognition: Dean's Honors List (Fall 2018, Fall 2020).

RESEARCH EXPERIENCE

WUHAN UNIVERSITY Wuhan, Hubei, China
Research Assistant - Advised By Qiming Liu Jan 2025–Jul 2025

- Performed XRD peak-profile/Rietveld fits to extract lattice parameters and peak positions; computed CV b-values via log i–log v regression and decomposed current as $i(V) = k_1 v + k_2 v^{1/2}$to estimate pseudocapacitance. Standardized and archived raw signals, derived metrics, and full metadata for downstream modeling.
- Developed a nonparametric baseline predictor (random forest / Gaussian process surrogate, per small-data best practices) for 100-cycle capacity retention using a compact feature set: electrochemical summaries (initial capacity, peak separation/polarization), XRD descriptors ((003)/(104) shift, c-axis spacing), and DFT-derived Na^+ migration barriers; validated with nested cross-validation and used predictive uncertainty to guide batch selection in the closed loop.
- Integrated group lasso as a first-pass screen with features grouped by modality (electrochemistry / structure / DFT) and by composition family, enforcing group-wise sparsity to filter irrelevant candidates and prioritize materials for synthesis and long-cycling tests.
- Ensured end-to-end reproducibility and data governance: maintained a version-controlled repo of Python/R scripts, enforced clear folder/schema and metadata standards, tracked datasets and experiment configs, locked environments (conda/renv), and generated auditable run logs—enabling deterministic reruns from raw files to figures.

PROFESSIONAL EXPERIENCE

HUBEI COMMUNICATION INVESTMENT GROUP CO., LTD. Wuhan, Hubei, China
Bridge Data Analyst, Department of Intelligence Analysis System May 2023–Sep 2023

- Built end-to-end QA/QC and normalization pipelines for multi-sensor, high-frequency SHM data (strain, acceleration, displacement, temperature, traffic), automating ingestion and cleaning in Python.
- Modeled and removed environmental/operational variability using regression analysis, producing temperature/traffic-corrected residuals that isolate structural signals for inference and monitoring.
- Conducted operational modal analysis (FFT/Welch, FDD/SSI) to estimate natural frequencies, mode shapes, and damping; tracked statistically significant shifts as stiffness-loss indicators.
- Developed moving-load analytics: influence-line estimation from strain/displacement + traffic, plus smoothing and wavelet decomposition to separate axle events from seasonal/thermal components.
- Applied unsupervised learning (PCA and clustering) to group sensors by structural behavior, reduce dimensionality, and detect redundant/rogue channels.
- Implemented data-driven alerting (EWMA/CUSUM and anomaly scores) with reproducible notebooks; validated findings with bridge engineers and inspections to close the loop from data to action.

UNITED WAY OF TOMPKINS COUNTY Ithaca, NY, United States
Data Analyst Intern Jan 2024-Jun 2024

- Led stakeholder-facing data collection and outreach: designed concise surveys and email copy, reconciled identities across CRM and partner files, and delivered clean, analysis-ready datasets to support multi-year reporting.
- Produced allocation and impact reporting—dashboards and presentations mapping inflows by source/area and outflows by program—and translated findings into clear, nontechnical updates for donors and partners.
- Built donor-engagement analytics (RFM segmentation; churn/return tracking) and wrote "next-best-action" briefs that guided fundraising cadence, messaging, and re-engagement strategy.
- Informed advertising strategy with data: identified high-yield programs/geographies, proposed targeting experiments (A/B tests), and set up recurring performance reports to iteratively refine campaigns.

INTERESTS

High-dimensional inference; sparse & structured regression (group/fused); robust estimation (Huber/quantile); fairness-aware modeling; uncertainty quantification; reproducible computational workflows for scientific discovery (materials & biostatistics).

TECHNICAL SKILLS

Programming: Python, R, MATLAB, C++
Statistical software: SAS
Data & tooling: SQL (MySQL), Git/GitHub, Quarto/R Markdown
Platforms: Linux (CentOS), macOS, Windows

LANGUAGES

Chinese: Fluent
English: Fluent