Ssaumya Jaiswal

ssaumyajaiswal07@gmail.com | State College, PA | https://github.com/saumya-k-j | https://github.com/saumya-k-j

Education

The Pennsylvania State University — Schreyer Honors College Aug. 2023 – Aug. 2026

B.S. in Data Science & Computational Mathematics Dean's List

Matthew Rosenshine Fund for Excellence — departmental award recognizing outstanding achievement in Statistics.

Honors Thesis: Learning Drift Functions and Deep Learning Approaches to the Poisson Equation in High-Dimensional Markov Chains.

Research Experience

Bayesian Statistics Research Assistant Feb. 2025 – Present

Penn State Eberly College of Science *State College, PA*

- Integrated probability theory, stochastic processes, and MCMC diagnostics to conduct advanced research on nonparametric variance estimation for reversible Markov chains.
- Leveraged Poisson equation formulations and Lyapunov drift functions to study long-run stability, ergodicity, and convergence rates of Markov chains under both geometric and subgeometric conditions.
- Developed and benchmarked deep-learning-based variance estimators against classical methods including batch means, spectral variance estimators, Metropolis–Hastings–based diagnostics, and importance sampling.
- Designed and ran high-dimensional simulation studies to evaluate robustness and efficiency of proposed estimators for slowly mixing and complex target distributions.

Machine Learning Research Assistant Apr. 2025 – Nov. 2025

Penn State College of Engineering *State College, PA*

- Co-developed MetaIoT, a novel meta-learning framework that enables intrusion detection models to rapidly self-adapt to unseen IoT networks—addressing one of the core unsolved challenges in real-world IoT security: extreme distribution shift.
- Implemented a bi-level optimization pipeline using MMD regularization to enforce domain alignment, ensuring feature representations remain stable across heterogeneous devices, protocols, and network conditions.
- Evaluated MetaIoT on IoT-23, ToN-IoT, Aposemat IoT, and UNSW-NB15, achieving high F1-scores, major reductions in false positives, and adaptation speeds surpassing conventional fine-tuning methods.
- Conducted robustness experiments measuring performance under unseen device types, mixed-protocol traffic, and novel attack families, demonstrating consistent generalization compared with static, non-adaptive intrusion detection models.
- Co-authored and presented the MetaIoT system in a full paper submitted to MLSys 2025.

Publications

Jaiswal, S. & Farooque, M. (2025). *METAIOT: A Cross-Dataset Meta-Learning System for Fast Adaptation in IoT Intrusion Detection.* Submitted to MLSys 2025.

Industry Experience

Data Analytics Intern May 2025 – Aug. 2025

TATA Power *Mumbai, India*

- Automated live SCADA data ingestion at one-second intervals using Python and SQL, eliminating manual processing and enabling high-frequency, real-time monitoring of grid operations.
- Designed and deployed an integrated Power BI analytics suite using DAX and M Query, providing engineers with interactive dashboards to analyze breaker tripping patterns, alarm frequency, and temporal trends.
- Reduced reporting and diagnostic latency by 26% by restructuring data flows and optimizing query performance within the analytics pipeline.
- Collaborated with electrical engineering and automation teams to build a full-stack web application (Python + JavaScript) for real-time visualization of operational metrics, improving situational awareness and enabling faster response to anomalies.

Data Analytics Intern Jun. 2023 – Aug. 2023

Variate Solar *Pune, India*

- Conducted data-driven feasibility analysis for transitioning large-scale industrial systems from conventional power to solar infrastructure using Python and R.
- Identified bottlenecks in modeling and simulation workflows and restructured pipelines, improving process efficiency by 25%.
- Developed SQL-based diagnostics and Power BI dashboards to monitor panel performance, energy yield, and cost-benefit metrics, reducing testing and implementation cycles by 30%.
- Supported engineering teams in interpreting analytical results and incorporating them into long-term renewable energy planning.

Teaching & Academic Leadership

Lead Academic Success Workshop Presenter Aug. 2024 – Present
Penn State Learning *State College, PA*

- Selected as one of Penn State's primary academic presenters, leading campus-wide workshops on metacognition, learning strategies, exam preparation, and time management.
- Designed research-informed curricula incorporating cognitive science principles such as spaced repetition, active recall, and self-explanation to improve student learning outcomes.
- Delivered high-impact sessions to 300+ students from diverse majors, helping them build sustainable study habits and stronger academic self-efficacy.

Mathematics Learning Assistant Jan. 2024 – May 2024
Penn State Department of Mathematics *State College, PA*

- Served as an instructional partner in a large-lecture Calculus II course (150+ students), reinforcing techniques in integration, series, and differential equations.
- Led weekly collaborative problem-solving workshops focused on conceptual understanding, multi-step reasoning, and exam-style practice.
- Communicated recurring misconceptions and difficulty patterns to instructors, helping refine lecture pacing and example selection.

Mathematics Tutor Aug. 2023 – Dec. 2023
Penn State Learning *State College, PA*

- Tutored students across algebra, trigonometry, Calculus I, and Calculus II, translating abstract concepts into intuitive explanations.
- Specialized in diagnosing foundational gaps and building personalized strategies to improve confidence and independent problem-solving.
- Facilitated small-group study sessions to encourage peer learning and collaborative reasoning.

Projects

LLM-Driven Portfolio Optimizer | *Python, LangChain, OpenAI API, FastAPI, Quantitative Finance* Apr. 2025 – Jul. 2025

- Designed a multi-agent LLM framework where specialized agents perform market sentiment analysis, technical forecasting, and risk assessment using real-time financial APIs.
- Integrated the Black–Litterman model, reinforcement-learning-guided rebalancing, and stochastic local optimization to generate adaptive, risk-aware portfolio allocations.
- Implemented a FastAPI backend to coordinate agent communication, execute optimization routines, and expose portfolio recommendations via REST endpoints.
- Built interactive dashboards to visualize agent reasoning, return trajectories, volatility shifts, and convergence behavior across simulated trading scenarios.

Google PageRank & Markov Chain Ranking Algorithms | *Python, NumPy, Markov Chains* Oct. 2024 – Dec. 2024

- Modeled the web as a directed graph and implemented ranking algorithms ranging from naïve inbound-link counting to a full Markov Chain–based PageRank system.
- Constructed transition matrices and analyzed convergence failures in reducible, periodic, and non-ergodic chains, motivating the need for damping and teleportation.
- Implemented PageRank with a random-surfer damping factor to ensure existence of a unique stationary distribution and stable convergence.

- Built counterexamples where structurally influential nodes receive higher PageRank scores despite having fewer direct links, illustrating the impact of global network structure.
- Benchmarked convergence times across randomly generated graphs and studied how connectivity, sink nodes, and spectral gaps influence power-iteration performance.

Machine Learning Algorithms From Scratch | *R, Statistical Learning Theory* Jan. 2025 – May 2025

- Derived core machine learning algorithms mathematically and implementing them entirely from scratch in R without ML libraries.
- Implemented k-Nearest Neighbors, linear regression, logistic regression, decision trees, custom distance metrics, loss functions, and gradient-based optimization procedures.
- Validated implementations by comparing predictions, residual patterns, and optimization behavior against R's built-in models, analyzing accuracy and numerical stability.
- Examined model assumptions, bias–variance tradeoffs, computational complexity, and failure modes across algorithms.
- Authored a technical report integrating derivations, code design, and empirical evaluation to demonstrate deep understanding of statistical learning foundations.

Advanced GUI Music Player | *Python, Tkinter, Pygame, Multithreading* Jan. 2025 – May 2025

- Developed a desktop music player with a responsive Tkinter UI supporting playlist creation, track navigation, and playback controls.
- Integrated Pygame for audio handling and implemented multithreading to separate UI rendering from audio playback, preventing interface freezing.
- Implemented metadata extraction (title, artist, duration) and optimized audio buffering to minimize latency when switching tracks.
- Structured the codebase using modular, event-driven design principles to improve maintainability and extensibility.

Technical Skills

Languages: Python, R, SQL, JavaScript, HTML/CSS, MATLAB
Frameworks: FastAPI, LangChain, Django, Tkinter, Pygame, SCADA systems
Libraries: NumPy, Pandas, TensorFlow, Scikit-learn, Matplotlib, OpenAI API, Power BI