Ssaumya Jaiswal

ssaumyajaiswal07@gmail.com | ssaumyakjaiswal.vercel.app/ | github.com/saumya-k-j | linkedin.com/in/ssaumyajaiswal

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   Mar 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   Sep 2025 – Present

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.

INDUSTRY EXPERIENCE

Data Analytics Intern   May 2024 – Aug. 2024

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.

TEACHING & ACADEMIC LEADERSHIP

Lead Academic Success Workshop Presenter   Feb 2025 – 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 & Math Tutor
Penn State Department of Mathematics & Penn State Learning

Jan 2024 – May 2025

State College, PA

- Served as a core instructional partner for a 150+ student Calculus II course, supporting faculty during lectures, reinforcing difficult topics, and providing real-time clarification during problem demonstrations.
- Delivered high-frequency tutoring across all 100-level mathematics courses—Algebra, Trigonometry, Calculus I, and Calculus II—using diagnostic questioning, targeted error analysis, and tailored explanations to address individual learning needs.
- Led structured, workshop-style problem-solving sessions each week, guiding students through multi-step reasoning, exam-level challenges, and collaborative exercises to strengthen conceptual understanding and mathematical maturity.
- Met regularly with course faculty to discuss student progress and learning barriers, presenting observations and recommending adjustments to pacing and instructional strategy to better support student comprehension.

PROJECTS

LLM-Driven Portfolio Optimizer | *Python,Graph Learning,Black–Litterman,FastAPI, Quantitative Finance*

Mar 2025 – Present

- Designed a multi-agent statistical forecasting framework where specialized models perform market sentiment extraction, short-horizon return estimation, and regime-shift detection using financial time-series and text-derived indicators.
- Integrated the Black–Litterman model with GNN-augmented covariance structure and a robustness-penalized SHLO + Hill-Climbing optimizer to generate adaptive, stability-aware portfolio allocations under uncertainty.
- Implemented a FastAPI backend to coordinate model interactions, compute Bayesian posterior returns, propagate graph-structured risk measures, and expose allocation recommendations via REST endpoints.

Google PageRank & Markov Chain Ranking Algorithms | *Python,NumPy, Markov Chains*

Oct 2025 – Present

- 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.
- Benchmarked convergence times across randomly generated graphs and studied how connectivity, sink nodes, and spectral gaps influence power-iteration performance.

Machine Learning Algorithms from First Principles | *R,Statistical Learning Theory*

Aug 2025 – Dec 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.

TECHNICAL SKILLS

Languages: Python, R, SQL, JavaScript, MATLAB, HTML/CSS
ML & Statistics: Deep Learning, Scikit-learn, TensorFlow, MCMC Diagnostics, Markov Chains, Bayesian Inference, MMD Domain Adaptation, Optimization Methods
Frameworks & Tools: FastAPI, Django, LangChain, Streamlit, Git, Jupyter, Power BI
Libraries: NumPy, Pandas, SciPy, Matplotlib, OpenAI API