Rohit Patra

rkumarpatra@gmail.com +1 718 679 5522 Website Scholar

Summary

Staff Research Scientist and former tenure-track statistics faculty member building production AI systems across LLM post-training, agent evaluation, RL/SFT training, reward design, contextual bandits, and uncertainty quantification. I develop statistically rigorous training, evaluation, and model-quality measurement methods that improve large-scale AI products.

LinkedIn, New York, NY August 2021 – Present

Staff ML Scientist, Applied Research

- *GenAI Modeling and Post-Training*: Core member of LinkedIn's GenAI modeling team, focused on task-specialized internal/on-prem models as lower-cost, governable alternatives to frontier models. Led SFT/RL post-training for production GenAI systems, spanning task-specific data construction, reward design, LLM-judge evaluation, offline validation, and launch criteria. In one production use case, shipped a specialized 3B on-prem model that replaced a frontier model API, improving quality by 14% while reducing serving cost by 30%.
- *Economic Graph Reasoning Agent (LLM+Graph)*: Core contributor to an LLM+graph reasoning agent that replaced semantic search with conversational search in a pilot over LinkedIn's heterogeneous economic graph; owned evaluation design, training-data pipelines, reward modeling, and metric definition.
- *Personalized Notifications (Decision Transformer RL)*: ML lead for a 10-engineer effort deploying a novel Decision Transformer + quantile-regression policy in a sequence-model-based, real-time notifications stack; owned architecture, training, and inference, delivering a 0.72% sessions lift.
- *Contextual Bandit Platform*: Architected and implemented a new bandit platform with efficient uncertainty quantification and coverage guarantees for LLM-scale models; built continual-retraining pipelines, policy-safety gates, interpretable evaluation metrics, and company-wide playbooks; delivered $100M+ incremental revenue and $+0.75\%$ sessions lift; led 12 engineers.
- *Data-Centric AI*: Proposed and built a weak-supervision platform and covariate/concept-shift detection system powering $10+$ models; reduced fake/abusive profiles by 15%; led 8 engineers.

University of Florida, Gainesville, FL

Tenure Track Assistant Professor, Department of Statistics September 2016 – August 2021

- NSF funded collaborations across ML, Economics, Astronomy, and Statistics; published in Journal of the Royal Statistical Society Series B, JASA, AoS, Bernoulli, Biometrics; co-PI on NSF DMS-2210662 ($200,000).

Education

Columbia University, New York, NY

- PhD in Statistics, Advisor: Bodhisattva Sen June 2016

Doctoral dissertation on using natural model constraints to do hyperparameter free inference and learning.

Indian Statistical Institute, Kolkata, India

- Bachelor's and Master's in Statistics July 2010

Selected Publications and Preprints

- Generative Sequential Notification Optimization via Multi-Objective Decision Transformers *In Review at KDD* (2025+) (with B. Ocejo and others)
- BanditLP: Large-Scale Stochastic Optimization for Personalized Recommendations *The ACM Web Conference 2026* (2026) (with P. Nguyen and others)
- Efficient Uncertainty Quantification with Guarantees via Sketching *In Preparation* (2025+) (with G. Dexter and J. McInerney)
- Computationally efficient Laplace approximations for neural networks. *NeurIPS Workshop on Bayesian Decision-Making and Uncertainty* (2024). (with S. Raha, K. Khare)
- Least squares estimation of a monotone quasiconvex regression function. *Journal of the Royal Statistical Society Series B* (2024). (with S. Mukherjee, A. L. Johnson, H. Morita)

— A nonparametric approach for estimating consumer valuation distributions using second-price auctions. Revise and resubmit, *Annals of Applied Statistics* (2023). (with S. Mukherjee, K. Khare)

— Dimension reduction for integrative survival analysis. *Biometrics* (2023). (with A. Molstad)

— On least squares estimation under heteroscedastic and heavy tailed errors. *Annals of Statistics* (2022). (with A. K. Kuchibhotla)

— Least squares estimation in a single index model with convex Lipschitz link. *Journal of the American Statistical Association* (2021). (with A. K. Kuchibhotla, B. Sen)

— Efficient estimation in single index models through smoothing splines. *Bernoulli* (2020). (with A. K. Kuchibhotla)

— A consistent bootstrap procedure for the maximum score estimator. *Journal of Econometrics* (2018). (with E. Seijo, B. Sen)

— Estimation of a two-component mixture model with applications to multiple testing. *Journal of the Royal Statistical Society Series B* (2016). (with B. Sen)

Selected Grants and Awards

— *NSF grant, Division of Mathematical Sciences – 2210662* ($200,000)
Title: Methodology for Qualitative Constraints in Semi-Parametric Models. Role: co-PI

— *National Fellowship in Basic Sciences* (Kishore Vaigyanik Protsahan Yojana), Department of Science and Technology, Government of India, 2005–2010.

Research Community Service

— Scientific Committee, IISA (2025); Co-organizer, UF Winter Statistics Workshop (2018, 2021); Co-organizer, Bayes Comp (2020).

— Graduate advising: committee member for 15+ theses.

— **Reviewer for**: JMLR, Journal of the Royal Statistical Society Series B, Journal of the American Statistical Association, AoS, AoAS, Econ Theory, Biometrika, and others.

Programming Skills

— PyTorch, TensorFlow, Python