

Narendra Mukherjee

✉ narendra.mukherjee@gmail.com
🌐 narendramukherjee.github.io
in [narendra-mukherjee](https://www.linkedin.com/in/narendra-mukherjee)

Employment

- July 2021- **Research Scientist**, *Philips Research*, Eindhoven, The Netherlands
- Probabilistic Machine Learning for marketing budget allocation and post-market surveillance in Philips' Personal Health business
 - Part of the AI, DS & Digital Twin research cluster
- July 2019-June 2021 **Machine Learning Scientist**, *TripAdvisor*, Needham, USA
- Bayesian and deep learning models of user-generated content and product recommendations for TripAdvisor's Experiences business
 - Worked alongside engineering to spearhead the adoption of a modern ML platform at TripAdvisor that can deploy containerized ML models and speed-up A/B testing

Education

- August 2019 **Ph.D. in Neuroscience and Quantitative Biology**, *Brandeis University*, Waltham, USA,
Dissertation title: Behaviorally relevant sensory cortical population dynamics in the rodent taste system
- **HHMI International Predoctoral Fellow** (<15% applicants selected internationally)
- May 2012 **Integrated BS-MS in Biological Sciences**, *Indian Institute of Science Education and Research*, Kolkata, India,
Dissertation title: Optimality and Courtship Behaviour in Zebrafish, *Danio Rerio*
- **Director's Gold Medal** (Best academic performance in a class of 80)

Ongoing research projects

Bayesian inference in dynamic models of online reputation systems, (*with Amin Rahimian and Shrabastee Banerjee*),

Using sequential neural posterior estimation (SNPE) for likelihood-free inference in simulations of user reviewing behavior on a digital platform

Bernoulli mixture Hidden Markov Models (BM-HMM) for large scale neural ensemble recordings, (*with Jian-You Lin*),

Variational inference in Bayesian HMMs with mixture emissions for robust modeling of massively high-dimensional time-series (like neural recordings) that are limited in size (by experimental constraints)

Bayesian nonparametric spectrum analysis, (*with Mark Goldstein*),

Variational inference in an Indian buffet process (IBP)-based spectral model with unknown number of sinusoidal components

Technical Expertise

Software Expert: Python, Unix/Linux, SQL(Hive, BigQuery, Postgres), LaTeX, HPC environments, Docker.

Intermediate: R, MATLAB, PySpark.

Working knowledge: C++, HTML, Kubernetes.

Modelling Machine Learning: Standard models for regression/classification, neural networks (deep networks, CNNs, RNNs, autoencoders), probabilistic graphical models (clustering, time-series models like HMMs, LDA, probabilistic PCA), Bayesian inference (including nonparametric priors with MCMC and variational-EM), NLP (TF-IDF, Doc2Vec, Word2Vec, ULMFiT, Transformers/BERT), Learning-to-rank (LambdaRank, LambdaMART).

Statistics: Frequentist techniques (parametric/non-parametric), Bayesian statistics (Hierarchical models, MCMC), computational neuroscience models (e.g. point-process models, drift-diffusion model of decision-making).

Frameworks: numpy, scipy, scikit-learn, Tensorflow/Keras/PyTorch/FastAI, PyMC3, Databricks, Spark, XGBoost, LightGBM.

Open-source projects ([Github](#))

- Hardware
 - Co-developed a Raspberry Pi-based hardware system to perform large-scale neural recordings in rodents.
 - Sampling rates of upto 40kHz from thousands of neural electrodes simultaneously.
 - Costs an order of magnitude less than any comparable commercially available solution.
 - Being used in 5 other neuroscience labs across the world - for details, please read our [Scipy 2017 paper](#).
- blech_clust
 - HDF5-based data management software to store, process and analyze neural voltage recordings upto several terabytes in size.
 - Tested on machines ranging from personal laptops to distributed clusters and cloud-computing environments.
 - Uses parallel computing to speed up the neural “*spike sorting*” pipeline by at least 20x.

Selected Publications (see expanded list on website)

- 2021 Banerjee S., Mukherjee N., Rahimian M. Amin. *Deep learning for simulation-based Bayesian inference of hidden parameters in online reputation systems. Workshop on Machine Learning for Consumers and Markets (MLCM) at the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD 2021)*
- 2019 Mukherjee N., Wachutka J., Katz D.B. *Impact of precisely-timed inhibition of gustatory cortex on taste behavior depends on single-trial ensemble dynamics. eLife. doi: doi.org/10.7554/eLife.45968.001*
- 2017 Mukherjee N., Wachutka J., Katz D.B. *Python meets systems neuroscience: affordable, scalable and open-source electrophysiology in awake, behaving rodents. Proceedings of the 16th Python in Science Conference. 97 - 104*
- 2016 Sadacca B.F., Mukherjee N., Vladusich T., Li J.X., Katz, D.B., Miller P. *The Behavioral Relevance of Cortical Neural Ensemble Responses Emerges Suddenly. Journal of Neuroscience. 36(3): 655 - 669*

Grants and Awards

- 2017-2019 \$29,513 (estimated) towards cloud computing resources on the Jetstream supercomputer of the XSEDE program of the National Science Foundation (NSF) (as administrator).
- 2014-2017 \$70,000 per year towards tuition and fellowship from the Howard Hughes Medical Institute (HHMI) as part of the International Predoctoral Fellowship.
- 2014 Pulin Sampat Memorial Award for the Best Teaching Fellow in the Life Sciences, Brandeis University.

Invited Talks

- 2021 **Bayesian imputation of missing feature values in product sort and recommendation at Tripadvisor.**
 - 14th ACM International Conference on Web Search and Data Mining (WSDM 2021)
- 2020 **When features go missing, Bayes’ comes to the rescue.**
 - PyData Global - [talk video](#)
- 2018 **Discrete cortical population activity states underlie taste processing and consumption behavior.**
 - Dept. of Mathematics and Statistics, Boston University, Boston, USA
 - Neuroscience Statistics Research Laboratory, Massachusetts Institute of Technology (MIT), Cambridge, USA
- 2017 **Systems neuroscience with Python: peering into the “black box”.**
 - Boston Python Meetup Group, Cambridge, USA
- 2017 **Building affordable, scalable and open-source tools in Python to study behaviorally relevant neural population dynamics.**
 - Center for Depression, Anxiety and Stress Research, McLean Hospital, Belmont, USA
 - Boston Python Meetup Group, Cambridge, USA

Personal

- Citizenship India
- Languages Hindi (native), Bengali (native), English (native/bilingual), Dutch (working proficiency: B1)
- Hobbies Long-distance road cycling, Travelling, Cooking