

Gautam Sabnis

ASSOCIATE COMPUTATIONAL SCIENTIST

600 Main Street, Bar Harbor, Maine 04609

850-320-4704 | gautam.sabnis@jax.org | gautam-sabnis | gautam-sabnis

Education

Florida State University

PHD IN STATISTICS

Tallahassee, FL, USA

August 2017

Florida State University

M.S. IN STATISTICS

- Specialization: Mathematical Statistics

Tallahassee, FL, USA

May 2014

Savitribai Phule Pune University

MASTER OF SCIENCE IN STATISTICS

Pune, MH, India

June 2011

St. Xavier's College (Autonomous)

BACHELOR OF SCIENCE IN STATISTICS

Mumbai, MH, India

June 2009

Experience

The Jackson Laboratory, Kumar Lab

ASSOCIATE COMPUTATIONAL SCIENTIST

Bar Harbor, ME, USA

Feb. 2023 - Present

- Machine Learning-Driven Analysis of Genetic and Phenotypic Correlates in Mouse Defecation Kinetics (In Preparation, 2025)*
 - Trained Bayesian Hierarchical (mixed effects) non-linear models to extract growth curve (kinetic) parameters from fecal boli patterns extracted using neural networks from thousands of animals across 62 strains in the open-field.
 - Applied a factor model to extract latent factors from a multi-dimensional behavioral feature space, distilling complex behavioral data into interpretable psychological constructs such as anxiety.
 - Performed genetic (heritability, genetic correlation, and association mapping) analyses across several fecal boli kinetic and open-field parameters to identify genetic correlates of anxiety.
 - Mentoring a USC student to carry out and extend the above analyses to 92 BXD strains.
- Visual detection of seizures in mice using supervised machine learning (Accepted in Cell Reports Methods, 2025)*
 - Trained a supervised latent variable model using regularized ordinal regression (specifically, cumulative link models) to predict univariate seizure intensity (no, low, medium, or high seizure groups) over an entire session.
 - Developed a temporally resolved seizure intensity scale by training an ordinal linear mixed model to predict seizure intensity from 1-minute bins of data, treating repeated measures as random effects and behaviors as fixed effects.
 - Performed rigorous cross-validation to estimate out-of-sample misclassification rates and validate performance on unseen data.
- Machine vision based frailty assessment for genetically diverse mice (GeroScience, 2025)*
 - Used advanced statistical machine learning models and tools to build predictive regression and classification models for predicting biological age for genetically diverse diversity outbred (DO) mouse population.
 - Created clearly documented code for reproducibility of results available as part of the publication.
- Highly accurate and precise determination of mouse mass using computer vision (Cell Patterns, 2024)*
 - Mentored Colby College student to carry out statistical modeling.

The Jackson Laboratory, Kumar Lab

BIostatistician II

Bar Harbor, ME, USA

Sep. 2019 - Jan. 2023

- A machine vision based frailty index for mice (Nature Aging, 2022)*
 - Used advanced statistical machine learning models and tools to build predictive regression and classification models for predicting the frailty score of a C57BL/6J or C57BL/6NJ mouse using behavioral features.
 - Used interpretable machine learning tools to provide biological interpretations for black box ML models.
 - Used quantile random forests to quantify the uncertainty in model predictions using prediction intervals.
 - Carried out extensive simulation studies requested during the peer-review process.
 - Created clearly documented code for reproducibility of results available as a part of the publication.
- Stride Level Analysis of mouse open field behavior using deep-learning-based pose estimation (Cell Reports, 2022)*
 - Used advanced hierarchical mixed-effects models and circular-linear regression models to characterize and validate gait and posture in autism spectrum disorder mouse models and models with gait disorder.
 - Used dimension reduction and clustering methods to identify different gait subtypes in 62 inbred strains.
 - Used classification algorithms to quantify the gait metrics' ability to distinguish mice across the C57 family.

The Jackson Laboratory, Kumar Lab

BIostatistician II

Bar Harbor, ME, USA

Sep. 2019 - Jan. 2023

- *Action detection using a neural network elucidates the genetics of mouse grooming behavior* (eLife, 2021)
 - Implemented Gaussian mixture models and k-means to identify grooming subtypes across several inbred strains using grooming metrics measured in an open-field assay.
 - Carried out statistical analysis to characterize the effects of sex and environmental factors on grooming behaviors in C57BL/6J and C57BL/6NJ.
- *A high-throughput preclinical assay for gait and posture deficits in the laboratory mouse using machine vision* (Preprint available upon request)
 - Performed univariate and multivariate statistical analyses of 234 mutant lines (2019 animals) to characterize gait and posture.
 - Identified 25% mutant lines that show altered gait and posture deficits in the KOMP2 phenotyping pipeline at the Jackson Laboratory.
- *A high-throughput machine learning based open field assay for nocifensive behavior in mice* (bioRxiv, 2023)
 - Implemented Bayesian ordinal logistic regression, using the probabilistic programming language Stan, to construct a nociception scale from behavioral video features.
 - Used statistical missing data techniques to impute missing gait metrics for certain strains.
- *Mouse Personalities - Behavioral Discrimination of Mouse Strains* (In preparation)
 - Used machine learning models to quantify the ability of different modalities of behavioral features (open field, grooming, gait etc.) to predict the strain of a mouse
 - Used Bulk RNA-seq data to predict behavioral traits. Used DESeq2 and WGCNA to identify differentially expressed genes and modules of co-expressed genes.

Dept. of Mathematics & Statistics, Boston University

POSTDOCTORAL RESEARCH ASSOCIATE

Boston, MA, USA

July 2018 - June 2019

- *Bayesian variable selection in linear regression models with instrumental variables* (arXiv, 2019)
 - Developed a novel semi-parametric quasi-Bayesian method for variable selection in high-dimensional instrumental variable linear models.
 - Implemented an efficient Markov Chain Monte Carlo sampler to sample from the quasi-posterior distribution.
 - Applied the method to calculate earnings of individuals in presence of endogenous variables from U.S census data.

Dept. of Statistics, University of Michigan

POSTDOCTORAL RESEARCH FELLOW

Ann Arbor, MI, USA

August 2017 - June 2018

- Explored a Bayesian logistic regression model for analyzing placement data on nearly 19,000 faculty in three disciplines.
- Designed a Metropolis-Hastings-within-Gibbs-sampler for making inferences on the placements network model in R.
- Implemented a divide-and-conquer approach in Bayesian factor models to characterize interactions in a high-dimensional **thyroid gene expression data** set in Matlab.
- Published a paper on compressed covariance matrix estimation with automated dimension learning for covariance structures arising via factor and PCA-type models.

Manuscripts

Machine vision based frailty assessment for genetically diverse mice

GeroScience, 2025

GAUTAM SABNIS, GARY A. CHURCHILL, VIVEK KUMAR

Visual detection of seizures in mice using supervised machine learning

Accepted in Cell Reports Methods, 2025

GAUTAM SABNIS, LEINANI HESSION, J. MATTHEW MAHONEY, ARIE MOBLEY, MARINA SANTOS, VIVEK KUMAR

Highly accurate and precise determination of mouse mass using computer vision

Cell Patterns, 2024

MALACHY GUZMAN, BRIAN Q GEUTHER, GAUTAM SABNIS, VIVEK KUMAR

A machine vision based frailty index for mice

Nature Aging, 2022

LEINANI E. HESSION*, GAUTAM SABNIS*, GARY A. CHURCHILL, VIVEK KUMAR (* EQUAL CONTRIBUTION)

Stride-level analysis of mouse open field behavior using deep-learning-based pose estimation

Cell Reports, 2022

KEITH SHEPPARD, JUSTIN GARDIN, GAUTAM SABNIS, ASAF PEER, MEGAN DARRELL, SEAN DEATS, BRIAN GEUTHER, CATHLEEN M. LUTZ, VIVEK KUMAR

Action detection using a neural network elucidates the genetics of mouse grooming behavior

eLife, 2021

BRIAN Q GEUTHER, ASAF PEER, HAO HE, GAUTAM SABNIS, VIVEK M PHILIP, VIVEK KUMAR

A high-throughput preclinical assay for gait and posture deficits in the laboratory mouse using machine vision

GAUTAM SABNIS, HAO HE, VIVEK KUMAR

Preprint available

A high-throughput machine learning based open field assay for nocifensive behavior in mice

GAUTAM SABNIS*, LEINANI E. HESSION*, KYUNGIN KIM, JACOB A. BEIERLE, VIVEK KUMAR (* EQUAL CONTRIBUTION)

bioRxiv, 2023

Bayesian Variable Selection in Linear Regression Models With Instrumental Variables

GAUTAM SABNIS, YVES ATCHADE, PROSPER DOVONON

arXiv, 2019

Compressed Covariance Matrix Estimation with Automated Dimension Learning

GAUTAM SABNIS, DEBDEEP PATI, ANIRBAN BHATTACHARYA

Sankhya Series A, 2019

A Divide and Conquer Strategy for High Dimensional Bayesian Factor Models

GAUTAM SABNIS, YOUNG LEE, DEBDEEP PATI, BARBARA ENGELHARDT, NATESH PILLAI

arXiv, 2016

Fusing surface and satellite-derived PM observations to determine the impact of international transport on coastal PM2.5 concentration in the western U.S.

NEHA BORA, TUO CHEN, DANA COCHRAN, KELLY DOUGAN, GAUTAM SABNIS, CHUANPING YU, ALEN ALEXANDERIAN, ARVIND SAIBABA, ELIZABETH MANNSHARDT, BRETT GRANT AND JESSICA MATTHEWS

Technical Report

Presentation

Machine vision based frailty assessment for genetically diverse mice

STATISTICAL METHODS, DEVELOPMENTS, AND CHALLENGES IN GEROSCIENCE

Indianapolis, USA

June 2025

A high-throughput preclinical assay for gait and posture deficits in the laboratory mouse using machine vision

CTC-RG 18TH ANNUAL MEETING

Bar Harbor, USA

September 2021

Bayesian Variable Selection in Linear Regression Models With Instrumental Variables

IISA 2019

Mumbai, India

December 2019

Bayesian Variable Selection in Linear Regression Models With Instrumental Variables

33RD NEW ENGLAND STATISTICS SYMPOSIUM

Hartford, USA

May 2019

A Divide and Conquer Approach to High-Dimensional Bayesian Factor Models

IISA 2018

Gainesville, USA

May 2018

A Divide and Conquer Approach to High-Dimensional Bayesian Factor Models

ENAR SPRING MEETING

Washington D.C., USA

March 2017

A Divide and Conquer Approach to High-Dimensional Bayesian Factor Models (3-minute speed talk)

FL ASA CHAPTER MEETING 2017

Jacksonville, USA

February 2017

Compressed Covariance Matrix Estimation with Automated Dimension Learning

IISA STUDENT PAPER COMPETITION 2016

Corvallis, USA

August 2016

Compressed Covariance Matrix Estimation with Automated Dimension Learning

ICSA APPLIED STATISTICS SYMPOSIUM

Atlanta, USA

July 2016

Skills

Technical	Statistical Methodology, Statistical Machine Learning, Multivariate Statistics, Scientific research & paper writing, Quantifying and drawing inferences from complex datasets
Programming Languages	R, Python, Julia, Stan, LaTeX, High-performance computing, Containers (Singularity), Git
Extracurricular	English, Hindi, Marathi
	6+ years training in Karate (Brown belt), 3+ years training in Tabla (Indian classical instrument. Cleared 3 theory & practical exams), Soccer (Regular starter on team for intramural soccer at Florida State University), Amateur Chess player

Awards

2025	Lightning Talk Winner , Statistical Methods, Developments, and Challenges in Geroscience	<i>Indianapolis, U.S.A</i>
2017	Yongyuan and Anna Li Award , Florida State University	<i>Tallahassee, U.S.A</i>
2017	Student Paper Award Winner , Statistics Inference for Biomedical Big Data, University of Florida	<i>Gainesville, U.S.A</i>
2017	Student Poster Competition Award , Florida Chapter of American Statistical Association	<i>Jacksonville, U.S.A</i>
2013	Best First Year Student Award in Computational Statistics , Florida State University	<i>Tallahassee, U.S.A</i>

Teaching

2019	Introduction to Statistics and Probability , Boston University	<i>Solo Instructor</i>
2018	Applied Statistics , Boston University	<i>Lab Instructor</i>
2018	Computational Methods , University of Michigan	<i>Solo Instructor</i>
2017	Introduction to Mathematical Statistics , Florida State University	<i>Solo Instructor</i>
2013 - 16	Introduction to Applied Statistics , Florida State University	<i>Solo Instructor</i>
2016	Applied Machine Learning , Florida State University	<i>Teaching Assistant</i>
2013	Fundamentals of Business Statistics , Florida State University	<i>Teaching Assistant</i>