Nikolai Vetr

Contact Information

Phone: (602) 578-9196
Email: nikolai.vetr@gmail.com
LinkedIn: linkedin.com/in/nikolai-vetr
GitHub: github.com/NikVetr

Education

PhD *2020*
University of California, Davis
GPA: 4.0/4.0, Anthropology, Population Biology, Data Science & Informatics

BA *2013*
Vanderbilt University
Earth & Environmental Sciences; Ecology, Evolution & Organismal Biology
GPA 3.9/4.0, Departmental Honors, *summa cum laude*

Projects

Dissertation Work

- Developed and implemented more efficient algorithms for computing phylogenetic likelihoods, mixing over high-dimension correlation matrices, and modularizing particular matrix operations under multivariate Brownian diffusion models of character evolution.
  - Reducing runtime for a typical analysis from weeks to hours
- Extended these models to discrete characters under the multivariate probit and explored their approximation in the context of truncated biogeographic dispersal.
- Applied these methods empirically to modern human, catarrhine, and fish datasets, also investigating their performance through extensive simulation experiments.

Recent Personal Projects

I've developed, fit, interpreted, described, and visualized results from:

- high-dimensional multilevel generalized linear mixed models (GLLMs) in Stan to heart transplant patient immune response data for time-series classification of rejection risk using non-HLA antigens straddling an acute rejection episode.
- multilevel GLMMs in Stan to consumer dietary and attitudinal response datasets measuring the effects of exposure to an advertisement advocating dietary change.
- a multilevel univariate Ornstein-Uhlenbeck model in Stan to the 'evolution' of nitrogen concentrations in manure ponds across California to predict values after an arbitrary amount of time has passed in an arbitrary pond.
- an efficient, conditional multivariate probit model in R that provides personalized movie rating predictions by exploiting basic properties of Schur complements.
- many dozens of minor scripts (< 100 LOC) devoted to data visualization, text mining, web scraping, replicating published analyses or algorithms from scratch, and exploring off-the-cuff ideas proposed by myself or colleagues.

Skills

Technical:

- Generalized Linear Models (e.g. spline, probit, Gaussian process, robust, etc.)
- Principle Components Analysis
- Missing Data Imputation
- Monte Carlo Methods
- Random Forests
- Diffusion Processes
- Support Vector Machines
- Neural Networks
- Model Comparison
- Expectation-Maximization
- Natural Language Processing
- Markov Chains (CTMC / DTMC)
- Measurement Error Modeling

Other: Science Communication, Teaching, Hitchhiking, Public Speaking, Backpacking, Data Visualization, Computer Assembly, Universal Design, Photography, Image Editing

Languages

Programming: R, RevBayes, Stan, BASH, Python, C++
Natural: Russian, English, Spanish

Leadership

Founder & Lead, Applied Bayesian Statistics Research Cluster *2019 - 2020*
Founded, coordinated, and supervised an interdisciplinary research cluster of 70+ scientists and statisticians across numerous career stages (PhD Student, Postdoc, PI, Industry Researcher), financially and spiritually sponsored by *Data Science & Informatics*.

Coordinator, Various Reading Groups *2014 - 2020*
Founded and coordinated reading groups on Python, Applied Linear Algebra, Bayesian Data Analysis, and Deep Learning, and frequently consulted on these and other topics for researchers both local and remote. Have also co-run reading groups on Computational Molecular Evolution, Paleoanthropology, Quantitative Genetics, and Machine Learning.

Teaching

Associate Instructor, UC-Davis *2015 - 2020*
Taught three quarters of an upper-division paleoanthropology course, two quarters of an upper-division evolutionary primatology course, and one quarter of an introductory course on human evolution. In these roles, I created or modified all the lab and lecture materials, designed and graded the assignments and tests, mentored students, delivered lectures, and supervised teaching assistants.

Teaching Assistant, UC-Davis & Vanderbilt University *2012 - 2020*
TA'd for the above, as well as the lab for an introductory cell biology course at Vanderbilt University. Duties here included designing and supervising labwork and lab-related assessments.

Outreach Lecturer, UC-Davis *2013 - 2020*
Gave multiple yearly talks on human evolution to elementary and middle school students on location and visiting UC-Davis, as well as during campus-wide events (e.g. *Picnic Day*). I've also volunteered for various workshops targeted at adults, e.g. to medical professionals on Natural Language Processing during UC-Davis Data Science Health Day.

Course Coordinator, Workshop in Applied Phylogenetics *2019*
Served as a course coordinator for a world-renowned, widely-attended, week-long workshop in applied computational Bayesian phylogenetics held in May 2019 at the Bodega Marine Laboratory.

Carpentries Instructor *2019*
Have completed Instructor Training and Checkout for The Carpentries organization and am fully certified to teach courses in the Data and Software Carpentries. Techniques learned here have been integrated into my own teaching on e.g. Bayesian statistics & phylogenetics.

Grants & Awards

Crook Travel Award ($500)	*2018*
1st Place Picnic Day Exhibit Award in "Secrets of Nature" Category	*2017*
Outstanding Graduate Student Teaching Award Nomination	*2016, 2019*
Summer Research Grant ($3k)	*2015*
NSF Graduate Research Fellowship ($152k)	*2015*
Summer Research Grant ($3k)	*2014*
Graduate Scholars Fellowship ($56k)	*2013*
Departmental Honors in Earth and Environmental Sciences	*2013*
Eugene H. Vaughan Undergraduate Research Assistantship in Geology ($13k)	*2012*
Geology Travel Grant ($500)	*2012*
Vanderbilt Undergraduate Summer Research Grant ($5k)	*2012*
Ross Family Scholarship ($60k)	*2012*
National Merit Scholarship ($20k)	*2009*