Nikolai Vetr

CONTACT INFORMATION
Phone: (602) 578-9196
Email: nikgvetr@stanford.edu
LinkedIn: linkedin.com/in/nikolai-vetr
GitHub: github.com/NikVetr/

Education

Postdoc, Stanford University *Current*
Computational Biology; Depts: Pathology, Genetics

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

BA, Vanderbilt University *2013*
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

- Generalized Linear Models (e.g. spline, probit, Gaussian process, robust, etc.)
- Principle Components Analysis
- Missing Data Imputation
- Monte Carlo Methods
- Random Forests
- Diffusion Processes / Time Series
- Support Vector Machines
- Model Comparison (e.g. via information theory, cross-validation, marginal likelihoods)
- Neural Networks
- Expectation-Maximization
- Clustering / Unclustering
- Markov Chains (CTMC / DTMC)
- Measurement Error Modeling
- Parallel & Distributed Computing

Languages

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

Leadership

Director, Board of Directors, Wild Animal Initiative *Current*
– I serve on the Board of Directors at the *Wild Animal Initiative* (WAI), a wild animal welfare non-profit research org. In addition, I serve as chair of WAI's Science and DEI Committees.

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 or co-founded groups on Python, Linear Algebra, Bayesian Data Analysis, Deep Learning, Computational Molecular Evolution, Quantitative Genetics, and Machine Learning.

Teaching

Associate Instructor, UC-Davis *2015 - 2020*
– Taught three terms of an upper-division paleoanthropology course, two terms of an upper-division evolutionary primatology course, and one term of an intro course on human evolution.
– Created or modified 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.
– Designed and supervised 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 and during campus-wide events (e.g. *Picnic Day*).
– 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 at the Bodega Marine Laboratory.

Carpentries Instructor, Data & Software Carpentries *2019*
– Completed Instructor Training and Checkout for The Carpentries organization. Integrated techniques learned into my own teaching on e.g. Bayesian statistics & phylogenetics.

Field & Labwork

Archaeological / Paleontological Excavator *2013*
Helped in summer excavation at La Ferrassie, a Neandertal fossil and artifact dig site.

Paleoecologist *2012 - 2013*
Examined how ecological and environmental conditions were recorded in enamel stable isotope ratios and dental microwear textures for marsupial taxa across Australasia.

Water Quality Analyst *2011*
Analyzed water quality across a large set of streams in the North Island of New Zealand.

Grants & Awards

Excellence in Data Science Community Training and Outreach ($1.5k) *2019, 2020*
Outstanding Graduate Student Teaching Award Nominee *2016, 2019, 2020*
1st Place Picnic Day Exhibit Award in "Secrets of Nature" Category *2017*
UC-Davis Summer Research Grant ($3k) *2014, 2015*
NSF Graduate Research Fellowship ($138k) *2015*
Graduate Scholars Fellowship ($56k) *2013*
Conference Travel Award ($500) *2012, 2018*
Eugene H. Vaughan Undergraduate Research Assistantship in Geology ($13k) *2012*
Vanderbilt Undergraduate Summer Research Grant ($5k) *2012*
Ross Family Scholarship ($60k) *2012*
National Merit Scholarship ($20k) *2009*