

Andrew Noble

Cambridge, MA

+1 (607) 351 6341

✉ andrew.e.noble@gmail.com

🌐 andrewenoble.com

github.com/andrewenoble linkedin.com/in/andrewenoble

Skills

Programming: Python, C/C++, Qt, SQL, Linux, Distributed Computing, Mathematica.

Tools: Scikit-learn, NLTK, Gensim, Matplotlib, Pandas, Numpy, Jupyter, Flask, D3.js, LaTeX.

Machine/Statistical learning: Logistic/Linear Regression, SVM, Neural Networks, K-Means, PCA, Hierarchical Bayes, NLP, Large-scale Monte Carlo Simulations, MLE, Bootstrap Resampling, Multivariate Stochastic Processes, Markov Chains, Undirected Graphical Models.

Experience

GNS Healthcare, Cambridge, MA, Data Scientist 2016–present

Insight Health Data Science, Boston, MA, Fellow 2016

- Consulted with Data Science team at Merck to prototype a web app, [HealthConnect.Online](#), for automating information sharing among patients and caregivers in online health communities.
- Cleaned 6,500 Reddit posts on Dementia and Alzheimer's forum using Jupyter, Pandas, RE, and NLTK Python packages for NLP. Stored posts in PostgreSQL database.
- Discovered a 3-topic Latent Dirichlet Allocation model, trained with Gensim, describing >90% of the unstructured data.
- Designed a web application with Bootcamp, D3.js, and Flask, served up by Unicorn on an AWS EC2 instance, that classifies user input and recommends related content.

University of California, Davis, CA, Principal Investigator 2013–2016

- Forged research team and led grant writing effort to win a \$600,000 interdisciplinary award from National Science Foundation with joint support from Directorates for Mathematical and Physical Sciences and Biological Sciences.
- Discovered a robust theoretical correspondence between biological oscillator networks and magnets with observational evidence found in a pistachio orchard.
- Developed [custom C/C++ software](#) to distribute Large-scale Monte Carlo Simulations over 192 cores of a Linux cluster. Bootstrapped statistical errors in over 100GB of synthetic data.
- Created animations of emergent fractal patterns in oscillator networks using Matplotlib.
- Co-organized two multidisciplinary working groups on Physics and Biology at Santa Fe Institute.

Universities of Texas, Maryland, and California, Postdoctoral Researcher 2008–2013

- Discovered new mathematical method for analyzing Multivariate Stochastic Processes enabling fast characterization of steady states describing evolutionary fixed points.
- Designed and deployed a plug-in menu of graph theory algorithms for QGIS, an Open Source Geographic Information System, using Qt, C++ Boost Graph Library, and PostgreSQL.

Cornell University, Graduate Student in Theoretical Physics 2002–2008

- Discovered potential for Higgs Boson self-coupling interactions at next-generation particle colliders to probe origins of matter/anti-matter asymmetry in universe.
- Built Mathematica notebooks to automate calculation of Feynman Diagram estimates for particle interactions and decay rates, simplifying thousands of pages of algebra into a single line and generating actionable predictions.

Education

Cornell University: PhD and MS in Theoretical Particle Physics and Cosmology 2008

Johns Hopkins University: MA in Secondary Science Teaching 2002

Carleton College: BA in Physics, Physics Department Honors 2000