

Daniel L. Klein

web: mesokurtosis.com • github: othercriteria • email: othercriteria@gmail.com • cell: 952.454.4197

Work experience

RealScout

October 2016 – November 2017

Data Scientist

- Led development of Recommended Listing feature: data exploration → agent-agent modeling → architecture plan → design mocks → user stories → implementation alongside engineers → deployment.
- Automated schema mapping for ETL of external listing databases (MLS) using a predictive model (gradient boosting) built on data and metadata features and trained on manually matched schemata.
- Answered general business intelligence and ops questions which required non-trivial amounts of SQL.
- Built out Bayesian models for sales price and days-on-market prediction with near-Zillow/Redfin accuracy.

Radius Intelligence, RealScout, Neighborly (rotations)

June – August 2016

Data Science apprentice (Catenus, 8VC)

TransForm Pharmaceuticals, Inc.

September 2006 – August 2008

Assistant Scientist, Scientific Computation

- Developed and deployed data integration software for laboratory automation platforms (HPLC, IR, spectrophotometer).
- Rapidly implemented ideas into software to support data analysis, molecular modeling, analytical chemistry method development, platform QA, etc.

Williams College, Dept. of Biology

June – August 2005

Research Assistant

- Planned and implemented experimental design and data analysis for field research project.

University of Minnesota, Dept. of Ecology and Evolutionary Biology

June – August 2003, 2004

Research Intern

- Developed and analyzed numerical results from novel model for ecological community assembly model.
- Participated in field work, data collection, and data entry.

Education

Brown University, Ph.D. Applied Mathematics

October 2016 (defended)

- Advised by Matt Harrison, with research focus on statistical inference in settings of extreme data sparsity or imbalance.
- TA for Intro Stats, Math Stats I/II, Recent Applications in Probability and Statistics.
- Organized and hosted speaker visits and meetings for Pattern Theory seminar.

Williams College, B.A. Mathematics and Biology

June, 2006

- Departmental honors in Biology, with thesis “Understanding aggregation in the membracid *Publilia concava*: using models to disentangle processes”.

Relevant coursework

- Foundational computer science, e.g., Data Structures, Design/Analysis of Algorithms, and Programming Languages.
- Bayesian Stats, Biostats, graduate Math Stats I/II, graduate Probability/Stochastics, Recent Applications in Probability and Statistics.
- Seminars in Graphical Models and in Bayesian Nonparameterics.
- Several courses in Computational Biology, covering use of dynamic programming and approximation algorithms to efficiently learn structural features from messy, incomplete, and/or inconsistent data.

Programming and technical comfort

- Enthusiastic: Python, R, C, SQL. Grudging: JavaScript, Matlab.
- Statistics, numerics, and visualization libraries: numpy/scipy/pandas/scikit-learn, TensorFlow, Stan, PyMC3, D3, etc.
- Familiarity with software engineering tooling (git, build systems, CI, etc.), web development (client-server, HTML, CSS, Haxyll, etc.), and cloud deployment (Docker, AWS, etc.).