
David J. Chudzicki

(518) 366 7303
69 Morisson Avenue
Somerville, MA 02144

dchudz@gmail.com
<http://davidchudzicki.com>
<https://github.com/dchudz>

WORK EXPERIENCE

Kaggle (7th employee)

Data Scientist & Software Engineer

January 2012 - present

- introduced optimization competitions at Kaggle, which have now become a staple of the site
- designed Kaggle's infrastructure for automated feature creation and model fitting/evaluating across a range of models and feature sets (for our oil & gas consulting arm of the business)
- spearheaded the use of interactive (Shiny) apps for exploratory visualization and visualizing the output of models. After some refinement, we used these apps in conversations with customers, leading customers to better understand and benefit from our work

Allstate Insurance

Predictive Modeler

October 2011 - January 2012

Associate Predictive Modeler

April 2010 - October 2011

- Built predictive models as the basis for Allstate's homeowners and auto insurance rating plans.
- Initiated and managed development of a tool to help the measure the effect of implementing a rating plan with deviations from the model.

EDUCATION

University of Chicago

M.S. in Mathematics

October 2007 - March 2010

- probability and statistics, functional analysis, stochastic differential equations, applied topology.
- Topic presentation: *Dynamical Systems: Bifurcation Under Symmetry*.

Swarthmore College

B.A. in Mathematics & Statistics, minor in Linguistics (GPA: 3.94)

September 2003 - June 2007

PROJECTS

- **Predcomps** (<http://www.davidchudzicki.com/predcomps/>): an R package for implementing Andrew Gelman's "average predictive comparisons", an approach for determining influence of input variables in a potentially complicated predictive model (e.g. without easily interpretable coefficients). Gelman mentioned the package [on his blog](#).
- **Simulated Knitting** (<http://blog.davidchudzicki.com/2011/11/simulated-knitting.html>): Represented knitted objects in Python, embedded these objects in 3D, and visualized them.

ACADEMIC RESEARCH

Department of Mathematics, University of Chicago

Graduate Research

April 2009 - March 2010

- Worked to apply differential geometry and statistics to problems in computer science. Investigated learning methods for high-dimensional (esp. semi-supervised) settings that exploit a lower-dimensional manifold structure by finding a good basis for functions which are smooth on the manifold.

Mathematics and Statistics Department, Swarthmore College

Research Assistant

June - August, 2007

- Assisted with the development and evaluation of estimators for the date of species' extinction given an incomplete fossil record. We used both analytical techniques and simulation.
 - Publication: S. Wang, D. Chudzicki, and P. Everson, 2009. "Optimal Estimators of the Position of a Mass Extinction When Recovery Potential is Uniform." *Paleobiology* 35:3, pp. 447-459.

TALKS

- **PyCon 2015, Montreal** (Tutorial): *Winning Machine Learning Competitions With Scikit-Learn*
- **Open Data Science Conference 2015, Boston**: *Data Workflows for Iteration, Collaboration, and Reproducibility*
 - A talk about using GNU Make to organize a reproducible data science pipeline. (slides).

PROGRAMMING

- **Languages**
 - Extensive: Python, R
 - Moderate: C#, JavaScript, Julia, MATLAB/Octave, SQL
- **Data visualization**: D3, ggplot2