William Levine

1710 Delaware St., Berkeley, CA   (202) 486-0972   wlevine@gmail.com   https://github.com/wlevine/

Experience

Research Assistant

Carnegie Mellon University   Sept. 2009 – Feb. 2016

- Isolated 200,000 particle physics collisions of interest from a dataset of 20 billion.
  - Without a labeled training set, used partial knowledge of the signal distribution and a modified nearest-neighbors algorithm for signal-background separation.
  - Worked as a member of CLAS, an international particle physics collaboration centered at Jefferson Lab particle accelerator.
- Wrote software and developed algorithm to cluster coincident detector signals in both simulated and experimental particle physics data.
  - Decreased rate of false cluster merges by 2 times and false cluster splits by 3 times.
  - This work became part of the standard analysis software used by dozens of members of the GlueX collaboration.

Student Developer

Ruby Science Foundation   Summer 2015

- Worked on NMatrix, a linear algebra library for Ruby, as part of Google Summer of Code.
- Wrote a second backend for advanced linear algebra features, allowing these features to be used with a generic LAPACK library rather than only with ATLAS.
- Decoupled the core NMatrix code from the more advanced features, which have complex external dependencies. This greatly simplified NMatrix installation.

Analyst

Students for Urban Data Studies   Spring 2016

- Worked with the Greater Hazelwood Community Collaborative, summarizing and presenting results from a community census they commissioned.
- Results available at http://www.greaterhazelwoodcensus.org/.

Projects

Writeups for projects available at https://wlevine.github.io/.

Personalized movie recommendations based on script text

- Supplemented a standard collaborative-filtering recommender system (based on R's `softImpute`) by inferring properties of the movie from its script.
- Used random forest regression and bag-of-words model on movie scripts to predict `softImpute` factors, allowing movies without ratings to be integrated into the recommender system.

Interlinear translations for foreign-language learners

- Wrote software that correlates sentences between an unstructured source text and a human translation of the source (with potential many-to-many correspondence between sentences). Used Bing machine translation API to identify matching sentences.
- Wrote scraper to retrieve different translations of the Bible from BibleGateway website.

Education

Carnegie Mellon University   2009 – 2016
PhD, Physics
M.S., Physics

University of Virginia   2005 – 2009
B.S., Physics

Skills

R, Python, pandas, numpy, scikit-learn, C++, Ruby, SQL, Git, Linux