Alan WILSON

(408) 242-5090
alan.w.wilson@gmail.com
https://github.com/aww
https://www.linkedin.com/in/alanwwilson/

Experience

2018–2023 **Lead Machine Learning Engineer**, *Glassdoor*, San Francisco, CA / Remote

- Modernized pay estimation with deep learning and organized builds with DVC
- Built and maintained a bespoke model for optimal pricing in the jobs marketplace
- Introduced structured model testing
- Built machine learning training and outreach for engineers and product managers

2015–2018 **Senior Data Scientist**, *Glassdoor*, San Francisco, CA

- Improved bot monitoring and flagging algorithms
- Built a job ad and direct response marketing demand forecasting system
- Built budget sizing and optimal pricing tools for job ad sales
- Ran an ecletic mix of small projects involving data processing and model building in offline python notebooks, hive, and some spark
- A/B testing platform development and maintainace (custom, in-house system):
  - built out new statisical components and added a completely new and much-refined UI
- Contributed many engagement and valuation insights around B2B products, CLV of marketing efforts, and user job search engagement

2014–2015 **Data Scientist**, *Glassdoor*, San Francisco, CA

- Brought to production core ML/data products at Glassdoor: **categorizing jobs**, and **salary prediction**
  1. End-to-end, from research and data exploration to model building and human testing to optimization for production.
- Early adopter of Tableau, company expert, author of a few very high-demand dashboards.

2014 **Fellow**, *Insight Data Science*, Mountain View, CA

- Developed the web app NewsSpectra.com which presents alternative news coverage of a topic on a spectrum of readability & detail.
- Full stack web dev. from Web scraping, NLP processing, and web app deployment with MySQL, AWS, Bootstrap, D3, etc.

2011–2013 **Post Doctoral Research Fellow**, *ATLAS Experiment*, Geneva, Switzerland

- Contributing to supersymmetry searches and the discovery of the Higgs boson
- Filtered massive datasets (> 1 TB) to find very rare signals using local and world-wide batch computing,
- Built C++ applications on top of the shared tools of a collaboration of 2000 scientists to calibrate, resolve ambiguities, and filter data.
- Developed a framework in Python and ROOT for specifying, building, and publishing plots to the web.
- Applied unit tests to publications by factoring out numerical quantities in JSON/YAML+Python.
- Controlled and monitored data acquisition, requiring quick reactions and efficient communication with colleagues.

2005–2010 **Graduate Student Research Assistant**, *ATLAS & DØ Experiments*, Michigan & Illinois

- Used a large stack of Monte Carlo simulations (from particle production to detector response), as well as extrapolations from data, to quantify signal and background.
- Increased signal to background separation in selection and identification problems using boosted decision trees.
- Computed 95% confidence intervals for new physical parameters using likelihoods built from data and statistical models of signal and background (with many nuisance parameters quantifying uncertainty).
- Managed Monte Carlo simulations by translating colleagues' informal requests into formal job specifications, testing and submitting the jobs, and monitoring the results; built tools in Python to streamline all of these steps.
- Contributed 100+ pages to documents describing the experiment's sensitivity to new physics.

2004–2005 **Research Assistant**, *ATLAS Experiment*, Univ. of Michigan, Geneva, Switzerland

- Validated software by broadly and systematically comparing alternative systems for unexpected discrepancies; found and reported on important bugs.
- Led a team of five students to complete assembly and testing of large detector components.

1994–2004 **Misc. research and teaching**, *U. of Washington & U. of Michigan*, Seattle & Ann Arbor

- Taught Calculus I & II, Differential Equations, etc. at the University of Michigan
- Teaching assistant for Discrete Math, Computer Graphics, and Digital Design at the University of Washington
- Developed software simulating X-ray imaging systems and for testing balloon and satellite experimental hardware in Geophysics at the U. of Washington

Education

2011 **Ph.D. Physics (experimental, high energy)**, *University of Michigan*, Ann Arbor

2003 **M.S. Mathematics**, *University of Michigan*, Ann Arbor

1999 **B.S. Computer Engineering & Mathematical Sciences (dual degree)**, *University of Washington*, Seattle

Proficiencies & technical interests

Almost every day

- Python, iPython & notebooks
- SQL (mostly for SQL Server, Hive, Presto, and SQLite)
- git, standard Linux/Unix tools (shell scripting, sed, emacs/vi, etc.)
- basic cloud computing (mostly AWS)
- numpy, scipy, matplotlib
- pytorch, tensorflow
- iPython notebook

Occasionally

- Airflow / Kubernettes
- MLFlow / DVC
- Statistical modeling
- A wide breadth of ML modeling techniques

Dabble in

- Hardware, IoT
- HTML, CSS, javascript, jQuery, web app deployment
- Web scraping
- Mathematica, Matlab, Octave
- Perl, Lisp dialects
- (Social) network analysis
- Coding and compression theory

Publications

ATLAS Collaboration. Measurements of four-lepton production at the z resonance in pp collisions at $\sqrt{s} = 7$ and 8 tev with atlas. *Phys. Rev. Lett.*, 112:231806, Jun 2014.

D0 Collaboration. Search for $z\gamma$ events with large missing transverse energy in $p\overline{p}$ collisions at $\sqrt{s} = 1.96\,\mathrm{TeV}$. *Phys. Rev. D*, 86:071701, Oct 2012.

ATLAS Collaboration. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29, 2012.

H J Yang, T Dai, A Wilson, Z Zhao, and B Zhou. A multivariate training technique with event reweighting. *Journal of Instrumentation*, 3(04):P04004, apr 2008.