

# Ryan Reece, Ph.D.

Data scientist / machine learning engineer / physicist

Mountain View, CA

ryan.reece@gmail.com

[rreece.github.io](https://github.com/rreece)

[linkedin.com/in/ryanreece](https://www.linkedin.com/in/ryanreece)

+1-817-888-7010

## EXPERIENCE

**Machine Learning Engineer** | Apr 2018 - Aug 2022 (4 yrs 4 mos)

[Cerebras Systems](#), Sunnyvale, CA

- Unicorn startup building high-performance machine learning accelerators, the first to achieve [Wafer Scale Integration](#)
- Developed high-performance input pipelines and reference model implementations in both [pytorch](#) and [tensorflow](#); helped maintain our model zoo
- Trained benchmark models and did exploratory optimization of various models for computer vision ([ResNets](#)) and NLP ([GNMT](#), [Transformer](#), [Linformer](#), [BERT](#), [RoBERTa](#))
- Helped develop a new normalization layer, [OnlineNorm](#), that uses streaming statistics to allow normalization of activations with small batch sizes [[arxiv:1905.05894](https://arxiv.org/abs/1905.05894)]
- Triage, explored, and tested customer-shared models; represented customer requirements to compiler engineers
- Assisted in customer engagement and support in meetings and on-sites, and the development of demos
- Co-authored a blog about [Getting started with PyTorch BERT models on the Cerebras CS-2 System](#)

**Artificial Intelligence Fellow** | Jan 2018 - Mar 2018 (2 mos)

[Insight Data Science](#), Palo Alto, CA

- 7 week fellowship: learned about data science and machine learning applications in a variety of business domains
- Developed cloud-based hyperparameter optimization platform: [HYPR.AI](#), for automating the testing of many ML models using AWS/Paperspace in docker containerized jobs

**Postdoctoral Research Fellow** | Jul 2013 - Aug 2017 (4 yrs 2 mos)

[Santa Cruz Institute for Particle Physics](#), The University of California, Santa Cruz, and  
The European Council for Nuclear Research ([CERN](#)), Geneva, Switzerland

- 10 years (postdoc and a Ph.D.) as a member of the ATLAS experiment, a 3000+ person collaboration looking for new physics in high energy proton-proton collisions at the Large Hadron Collider (LHC)
- Long involvement in codebase of more than 10 million lines of C++ and almost as many lines of Python
- [Expert in petabyte data reduction](#) (ATLAS ~10 PB/year), world-wide grid computing, and [data visualization](#) as a user and primary supporter of our group's 200-CPU computing cluster, accumulated more than [350k CPU-hours](#)
- Lead analysis groups as "Editor" in different searches for signals of supersymmetry and exotic decays, contributed to 6 research publications, and defended their approval
- 2015-17, full-time support the operations of the [data acquisition system \(DAQ\)](#) and detector monitoring systems of the SCT (a tracking sub-detector in ATLAS)
- 2016-17, built more [expertise in machine learning techniques](#), deep learning frameworks using Keras to build CNNs for [particle classification](#), and another project using sklearn for [anomaly detection](#)

**Graduate Researcher** | Jun 2006 - Jul 2013 (7 yrs)

[The University of Pennsylvania](#), Philadelphia, PA, and

The European Council for Nuclear Research ([CERN](#)), Geneva, Switzerland

- First summers as a student with Penn (2006-08) at CERN participating in the [integration and commissioning of custom electronics](#) for the Transition Radiation Tracker (TRT), the outermost sub-detector of the ATLAS tracker
- 2009-12, throughout most of the running of the LHC, rotated the [on-call responsibility](#) for the TRT DAQ

- Ph.D. research with the data from ATLAS focused on the identification of decays of tau leptons and their use in searches for new physics, a [pattern recognition](#) problem to identify a type of particle
- 2009-10, was the lead developer of the cut-based tau identification used with the first ATLAS data
- 2010-12, helped develop advanced tau identification using [Boosted Decision Trees \(BDTs\)](#) which superseded the above
- Knack for developing data analysis frameworks: e.g. [pyframe](#) has been used by several analyses in ATLAS
- The ATLAS and CMS experiments at the LHC [discovered the long-sought-after Higgs boson](#), evidence of which was announced on July 4, 2012 [[arxiv:1207.7214](#)]

## EDUCATION

- **Ph.D. Experimental Particle Physics**, The University of Pennsylvania (Philadelphia, PA), Jun 2006 - Jul 2013  
thesis: “[A search for new physics in high-mass ditau events in the ATLAS detector](#)”
- **B.S. Physics with Honors**, The University of Texas (Austin, TX), Aug 2003 - May 2006  
thesis: “Late pulsing in the Hamamatsu R1408 PMT used in the Sudbury Neutrino Observatory”

## PUBLICATIONS

- Chiley, V. *et al.* (2019). Online Normalization for Training Neural Networks. [arxiv:1905.05894](#).
- Albertsson, K. *et al.* (2018). Machine Learning in High Energy Physics Community White Paper. [arxiv:1807.02876](#).
- As a member of the ATLAS collaboration since June 1, 2008, I am an “author” of more than 800 publications ([google scholar](#), [inspire](#)), however, my list of selected publications is here: [reece.github.io/publications](#), but in particular:
  1. Search for supersymmetry in a final state containing two photons and missing transverse momentum in  $\sqrt{s} = 13$  TeV  $pp$  collisions at the LHC using the ATLAS detector. *European Physical Journal C*, 76, 517 (2016). [arxiv:1606.09150](#).
  2. Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in  $pp$  collisions at  $\sqrt{s} = 8$  TeV. *European Physical Journal C*, 75, 303 (2015). [arxiv:1412.7086](#).
  3. A search for high-mass resonances decaying to  $\tau^+\tau^-$  with the ATLAS detector. *Physics Letters B*, 719, 242-260 (2013). [arxiv:1210.6604](#).
  4. Performance of the ATLAS detector using first collision data. *Journal of High Energy Physics*, 9, 56 (2010). [arxiv:1005.5254](#).

## SKILLS

- **General:** deep learning, statistical analysis, data visualization, data-driven modeling, anomaly detection, neural network classifiers, boosted decision trees, petabyte data reduction, object-oriented design, polymorphic interfaces, writing technical reports, working independently and in groups, presenting my ideas, graduate level physics and mathematics
- **Programming languages (fluent):** C/C++/STL (17+ years), Python (15+ years);  
(experienced): javascript, SQL; **Markup languages:**  $\text{\LaTeX}$ , Markdown, (x)html with css
- **Data science software:** [pytorch](#), [tensorflow](#), [keras](#), [HuggingFace](#), matplotlib, numpy, scipy, scikit-learn, pandas, jupyter, AWS (EC2, S3), docker, singularity, ROOT, RooStats, TMVA
- **General software:** Linux, bash, git, svn, UML, QT, Mathematica
- **Hobbies:** poker, philosophy, cycling, climbing