Ryan Reece, Ph.D.

ryan.reece@gmail.com
rreece.github.io
linkedin.com/in/ryanreece
+1-817-888-7010

Data scientist / machine learning engineer / physicist
Mountain View, CA

EXPERIENCE

Machine Learning Engineer | Apr 2018 - present
Cerebras Systems, Los Altos, CA

- Stealth-mode startup building high-performance machine learning accelerators
- Developed high-performance, containerized tensorflow input pipelines with tf.data
- Trained benchmark models and did exploratory optimization of various computer vision (ResNets) and NLP models (seq2seq, GNMT)

Insight Artificial Intelligence Fellow | Jan 2018 - Mar 2018
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
Santa Cruz Institute for Particle Physics, The University of California, Santa Cruz, and
The European Organization 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 two searches for signals of supersymmetry and other exotic decays, managed a team of 5-6 analysts, eventually wrote the research paper, and defended its approval—a period of about a year
- 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)
- 2017, built more expertise in machine learning techniques, deep learning frameworks, and probabilistic databases: using Keras to build CNNs for particle classification, and another project using BayesDB and sklearn for anomaly detection

Graduate Researcher | Jun 2006 - Jul 2013
The University of Pennsylvania, Philadelphia, PA, and
The European Organization 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

EDUCATION

- **Ph.D. Experimental Particle Physics**, The University of Pennsylvania (Philadelpha, 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"

SKILLS

- **General:** 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 (11+ years);
  (experienced): javascript, SQL; **Markup languages:** LaTeX, Markdown, (x)html with css
- **Data science software:** matplotlib, numpy, scipy, scikit-learn, pandas, jupyter, Keras, tensorflow, BayesDB, AWS (EC2, S3), docker, ROOT, RooStats, TMVA
- **General software:** Linux (Redhat/SLC/Ubuntu/Debian), bash, git, svn, UML, QT, Mathematica
- **Hobbies:** climbing, philosophy, poker, running