Toronto, ON, Canada
github.com/yifever

# Yifei Han

Data Scientist

Google Scholar Profile
linkedin.com/in/yi-fei-han
yhan@physics.utoronto.ca

## Skills

| | |
|---|---|
| **Tools and Languages** | C++, Python, Git CI/CD, LaTeX, SQL, ROOT (R-like C++ library) |
| **Distributed Computing** | Linux, CERN grid, AWS, Docker |
| **Productivity** | Agile Development, Jira, Slack |
| **Quantitative Research** | Statistics/Modeling, Tensorflow, Keras, XGBoost |
| **Communication** | English (native), Chinese Mandarin (native), Japanese (JLPT N2) |

## Experience

**Researcher - Data Science** **September 2017 - Current**
*ATLAS Experiment | CERN* *Toronto, Canada*
*ATLAS is a general-purpose particle physics experiment at the Large Hadron Collider at CERN designed to answer questions about the fundamental forces of nature. As of 2022, it has over 5000 collaborating members from 181 institutions in 42 countries.*

- Led efforts in a team of 10 to improve data quality and increased the statistical significance from 99.4% to over 99.9999999% ($4\sigma$ to $6.5\sigma$) for a signal model of less than $e^{-13}$ effect
- Manipulated 1000+ TB datasets use over the CERN computing grid (similar to AWS), filtering down to a 2 GB final dataset for causal inference, which was done with a statistical model of 500+ sources of error all constrained with data
- Performed A/B testing comparing deep neural networks built with TensorFlow, decision trees built with XGBoost, and a strong baseline of physics logic filters, taking into account signal significance related KPIs and business logic (physics), increasing signal acceptance by a factor of 3
- Wrote and edited over 500 pages of publication ready presentations and documents, some of which have 100+ citations, and presented at conferences both internal and external

**Developer - Detector Simulation** **August 2017 - September 2021**
*ATLAS Experiment | CERN* *Toronto, Canada*

- Contributed to an open-source physics engine (acts) to simulate detector response to charged particles
- Optimized physics simulation time with a library of pre-processed detector responses as a function of input parameters using C++ and python, for a simulation engine used by 2000+ scientists
- Improved simulation time by a factor of 10, from 28.1s/event to 2.82s/event

**Teaching Assistant** **September 2016 – September 2021**
*University of Toronto* *Toronto, Canada*

- Ran weekly tutorials, hosted office hours, and graded exams for undergraduate physics courses of up to 150 students
- Communicated topics include mechanics, electromagnetism, quantum mechanics, and relativity to non-specialists

## Education

**PhD in Physics**, *University of Toronto* September 2016 - 2022
**Bachelor of Science in Physics and Mathematics**, *University of California, Los Angeles* September 2012 - June 2016

## Select Publications

*As an ATLAS author, I have my name on every published ATLAS paper. I contributed significantly to these following publications:*
ATLAS Collaboration, *Measurement of fiducial and differential $W^+W^-$ production cross-sections at $\sqrt{s} = 13$ TeV with the ATLAS detector.* The European Physical Journal C, 79(10), Oct 2019. arXiv:1905.04242[hep-ex]
ATLAS Collaboration, *Search for electroweak diboson production in association with a high-mass dijet system in semileptonic final states in $pp$ collisions at $\sqrt{s} = 13\mathrm{TeV}$ with the ATLAS detector.* Phys. Rev. D 100, 032007. arXiv:1905.04242 [hep-ex]
ATLAS Collaboration, *The Fast Simulation Chain in the ATLAS experiment.* EPJ Web Conf. 251 03012 (2021). DOI: 10.1051/epjconf/202125103012

## Hobbies

- Hosted online table-top RPG platform using AWS for friends using open source software and node.Js/JavaScript