

# Danny Luo

Xiao Luo | [luoxiao96@hotmail.com](mailto:luoxiao96@hotmail.com) | [dluo.me](http://dluo.me) | [github.com/PiercingDan](https://github.com/PiercingDan) | [linkedin.com/in/dannyluo](https://linkedin.com/in/dannyluo)

Data scientist with strong mathematical background and experience in big data, machine learning, statistics and academic research. Educator of mathematics and data science.

## Skills & Experience

<b>Technical Skills</b>	<i>ML Programming Big Data</i>	Neural networks, gradient boosted models (XGBoost) Python (scikit-learn, numpy, pandas, tensorflow), Linux, Scala, Java, R, SQL Apache Spark (Python, Scala), Hadoop, AWS (EC2, S3), Azure, Databricks
<b>Projects</b>	<b>Residential Real Estate Valuation</b> — ProvidentAI	<ul style="list-style-type: none"><li>◆ Achieved 15% prediction error on housing prices using XGBoost in Python on EC2</li><li>◆ Extracted, cleaned and pre-processed over 13 million records from remote SQL database in Python and SQL</li></ul>
	<b>Optimal Digital Map Placement in Toronto</b> — 3rd place project at HackOn(Data) 2016	<ul style="list-style-type: none"><li>◆ Pinpointed and ranked 200+ hotspots in Toronto for optimal map placement by employing unsupervised learning in Python on transit, cultural centre and social media data</li><li>◆ Mined 8000 check-ins to improve model using Foursquare API, Spark on Databricks</li></ul>
	<b>Guide: Spark with Jupyter on AWS</b> — 200+ Github stars	<ul style="list-style-type: none"><li>◆ Created guide on launching Spark clusters on EC2 with Jupyter and S3 integration</li><li>◆ Presented to 100+ people under invitation at Toronto Apache Spark Meetup</li></ul>
<b>Groups</b>	<b>University of Toronto Data Science Team</b>	<ul style="list-style-type: none"><li>➔ Parallelized CPU processing of kaggle CT images in Python on custom ML infrastructure (EC2, S3), resulting in 10x speed improvement, 8x disk size reduction</li><li>➔ Organized and mentored hands-on data science workshops for 30+ UofT students</li></ul>
	<b>HackOn(Data) 2017 Mentor &amp; Judge</b>	<ul style="list-style-type: none"><li>➔ Mentored 1st and 2nd place teams, judged data science projects of 15+ teams</li></ul>
<b>Other</b>	<b>Neural Networks:</b> Built low-level CNN, RNN, GAN in Python (numpy) from Stanford cs231n	
	<b>Coursera Algorithms I, Princeton:</b> Completed course, implemented algorithms in Java	

## Education

2014 - present **University of Toronto, Candidate for Bachelor of Science** — Graduating in May 2018  
Mathematics and Physics - CGPA: 3.97, Dean's List 2015/16/17

## Employment

5/2017 - 8/2017	<b>Scientist, Co-op</b> — <i>ZeroGravityLabs, LoyaltyOne</i>	<ul style="list-style-type: none"><li>• Formulated decentralized loyalty programs on Blockchain, presented to CTO &amp; other senior executives, leading to creation of strategic taskforce based on this work</li><li>• Automated Ethereum blockchain deployment on EC2 using bash, go-ethereum, AWS CLI</li><li>• Processed 2 TB of transactional data for economic simulation in Spark (Scala) on Azure</li></ul>
2015 - present	<b>Teaching Assistant</b> — <i>UToronto Dept. of Mathematics</i> — 200+ students over 6 courses	<ul style="list-style-type: none"><li>• Developed content and curriculum for MAT245: Mathematical Methods in Data Science</li></ul>
5/2016 - 8/2016	<b>Particle Physics Undergraduate Researcher</b> — <i>UToronto Dept. of Physics, NSERC USRA</i>	<ul style="list-style-type: none"><li>• Analyzed CERN LHC data and Monte Carlo simulations for Higgs Boson in Python, C++</li></ul>
5/2015 - 8/2015	<b>Astrophysics Undergraduate Researcher</b> — <i>UWaterloo Dept. of Astronomy, NSERC USRA</i>	<ul style="list-style-type: none"><li>• Mapped evolution of galaxies through analysis of 10+ datasets in Python</li></ul>

## Awards

2016/15 **NSERC Undergraduate Student Research Awards:** National award for scientific research  
2015 **University of Toronto Scholar:** Top 150 students per year of study (10000+ students)

## Interests

**Leisure** Salsa, Bachata, Volleyball, Fitness, Piano, Guitar, Jazz, Poetry, Literature, Creative Writing  
**Academic** Cosmology, Differential Geometry, Probability, Topological Data Analysis, Manifold Learning