

# Ilia Sucholutsky

(416)-543-1369

[isucholu@uwaterloo.ca](mailto:isucholu@uwaterloo.ca)

[ilia10000.github.io](https://ilia10000.github.io)

## Education

**Candidate for PhD, Statistics, University of Waterloo** (Sept. 2017 – May 2021)

Supervisor: Dr. Matthias Schonlau

**BMath with Distinction, Statistics, University of Waterloo** (Sept. 2014 – Aug. 2017)

## Research Experience

**PhD Research, University of Waterloo** (Sept. 2017 - Present)

Topic 1: Deep learning for anomaly detection and loss restoration

Topic 2: Soft-label dataset distillation

Topic 3: "Less than one"-shot learning: Learning N classes using  $M < N$  samples

Thesis: Learning with almost no data

Supervisor: Dr. Matthias Schonlau, Statistics and Actuarial Science

**Research Associate, University of Waterloo** (May 2017 - Aug. 2017)

Topic: Deep learning for anomaly detection in safety-critical real-time embedded systems

Supervisor: Prof. Sebastian Fischmeister, ECE

**Research Assistant, University of Waterloo** (May 2017 - Aug. 2017)

Topic: Customizable neural networks in Stata

Supervisor: Dr. Matthias Schonlau, Statistics and Actuarial Science

**Research Assistant, University of Waterloo** (Nov. 2016 - Apr. 2017)

Topic: Text mining with n-gram variables

Supervisor: Dr. Matthias Schonlau, Statistics and Actuarial Science

## Industry Experience

**Vice President Research, StratumAI** (Aug 2020 - Present)

Leading research and development of algorithms to make mining efficient and environmentally sustainable

Adapting deep learning methods for 3D spatial regression on highly sparse datasets

Improving explainability and uncertainty estimation of neural network predictions

Developing methods for integrating geological insights as priors into neural network training

**Data Guru (Research Team), Athos** (May 2016 – Aug. 2016)

Created Luigi pipelines to automate away EMR jobs saving each team member an average of 8-10 hours per week

Increased calibration accuracy by 90% by creating patch to recalculate values in backend and push to users' app

Developed and automated system for surfacing defect analytics internally, leading to 10% increase in contact quality

Investigated root cause of churn and improved user experience through visualizations of progress metrics

Extracted new features from multi-channel time-series data including real 3D motion and muscle-use timing

Performed study on data from elite and non-elite athletes to determine what leads to high performance

**Data Scientist, Capital One** (May 2015 – Aug. 2015)

Completed cross-functional Facebook advertising insourcing project start-to-finish:

Increased app volume to 260% and decreased cost per app by 50% with net benefit of \$2 million per year

Interfaced with Facebook API programmatically to create a fully automated data pipeline

Used NLP techniques to identify relevant demographic segments

Performed rank-order and slope analysis to identify optimal bids for demographic segments

Transitioned model and data infrastructure from Linux server to Hadoop

## Tech Stack

Languages: Python, R, C++, C, SQL

Databases: Teradata, MySQL, MongoDB

Servers: Linux, Hadoop, AWS, Shell Scripting, \*NIX

Tools: Pytorch, Keras, TensorFlow, Numpy, Pandas, scikit-learn, Scipy, matplotlib, Git

## Consulting Experience

### **Consulting and Advising**

**(Jan. 2017 – Present)**

Helped rapidly growing accounting firm automate data-entry/bookkeeping pipelines  
Received controlled goods clearance and provided ML consulting to a large defence contractor  
Worked with a mining startup to improve their results when using deep learning to model extremely sparse 3D data  
Consulted fintech startup on improving data-efficiency using latest few-shot, active, and online learning research  
Developed R&D plan for deepfake entertainment startup (10M+ users) to improve models and enable efficient scaling  
Advised (pro-bono) early-stage startups on developing ML systems, planning ML R&D, getting the most out of data, etc.

### **President, UW Apprentice**

**(Jan. 2015 – Jan. 2017)**

Managed diverse team of 40 students spanning two universities  
Provided pro-bono consulting to 20+ startups ranging from pre-seed to Series B  
Negotiated sponsorship deals with firms including Blackberry, Kik, Capital One, and Velocity to fund this initiative

## Academic and Administrative Experience

### **Instructor, University of Waterloo**

**(Jan 2020 - April 2020)**

Course: Stat 231 - Statistics  
Evaluation: Received a 4.4 (out of 5) weighted score on student evaluations

### **Co-founder & Treasurer, Stats Anti-Depression (S.A.D.) Club**

**(Jan 2018 - Present)**

Started grassroots mental health initiative for graduate students  
S.A.D. Club provides students with a safe, relaxed, social setting where they can unwind, socialize with peers, have some nutritious snacks, and hopefully just overall reduce their stress levels

### **Teaching Assistant, University of Waterloo**

**(Jan. 2017 – Apr. 2017)**

Course: Math 135 - Algebra for Honours Mathematics

### **Teaching Assistant, University of Waterloo**

**(Sept. 2016 – Dec. 2016)**

Course: Math 114 - Linear Algebra for Science

### **Student Councillor, FEDS, University of Waterloo**

**(May 2016 – Apr. 2017)**

Elected to represent undergraduate mathematics students

## Research

### **Peer-reviewed**

Sucholutsky, I., Schonlau, M. Soft-label dataset distillation and text dataset distillation. Preprint at <https://arxiv.org/abs/1910.02551v2>. Accepted for publication in the proceedings of IJCNN 2021.

Sucholutsky, I., Kim, N., Schonlau, M. One line to rule them all: generating LO-shot soft-label prototypes for kNN. Preprint at <https://arxiv.org/abs/2102.07834>. Accepted for publication in the proceedings of IJCNN 2021.

Sucholutsky, I., Schonlau, M. "Less than one"-shot learning: learning N classes using  $M < N$  samples. Preprint at <https://arxiv.org/abs/2009.08449>. Accepted for publication in the proceedings of AAAI 2021. See below for press coverage on this research.

Sucholutsky, I., Schonlau, M. SecDD: efficient and secure method for remotely training neural networks (Student Abstract). Preprint at <https://arxiv.org/abs/2009.09155>. Accepted for publication in the proceedings of AAAI 2021 Student Abstract and Poster Program.

Sucholutsky I, Schonlau M. 2021. Optimal 1-NN prototypes for pathological geometries. PeerJ Computer Science 7:e464 <https://doi.org/10.7717/peerj-cs.464>

Sucholutsky, I., Narayan, A., Schonlau, M., Fischmeister, S. Pay attention and you won't lose it: a deep learning approach to sequence imputation. PeerJ Computer Science. 2019.

Sucholutsky, I., Narayan, A., Schonlau, M., Fischmeister, S. Deep learning for system trace restoration. In the proceedings of the International Joint Conference on Neural Networks 2019 (IJCNN 2019).

Sucholutsky, I., Schonlau, M. ConvART: Improving adaptive resonance theory for unsupervised image clustering. Journal of Computational Vision and Imaging Systems. Dec 2018, 4(1)

Schonlau, M., Guenther, N. Sucholutsky, I. Text mining using ngram variables. The Stata Journal. Dec 2017, 17(4), 866-881. Preprint also available at <http://ssrn.com/abstract=2759033>

### ***Under review or in progress***

Sucholutsky, I., Wong, A., Schonlau, M. Probing the limits of few-shot learning with deep neural networks. In progress.

Sucholutsky, I., Schonlau, M. Better label initialisations for soft-label dataset distillation. In progress.

Sucholutsky, I., Narayan, A., Schonlau, M., Fischmeister, S. Search and rescue: a neural architecture search approach to sequence imputation. In progress.

### ***Presentations***

AAAI 2021 Main Track - "Less than one"-shot learning: learning N classes using M<N samples.

AAAI 2021 Student Abstract and Poster Program - SecDD: Efficient and Secure Method for Remotely Training Neural Networks

International Joint Conference on Neural Networks 2019 (IJCNN 2019) - Deep Learning for System Trace Restoration

4th Annual Conference on Vision and Intelligent Systems (CVIS 2018) - ConvART: Improving Adaptive Resonance Theory for Unsupervised Image Clustering

University of Toronto StartAI 2018 Conference - Making the Most of Graduate Research in AI

Statistics Society of Canada Annual Meeting 2018 - Deep Learning for Lost Data Restoration and Imputation. Abstract available at <https://ssc.ca/en/meeting/annual/presentation/deep-learning-lost-data-restoration-and-imputation>

University of Waterloo Data Science Club - Part 1: Deep Learning for Lost Data Restoration and Imputation

University of Waterloo Data Science Club - Part 2: Breaking into Deep Learning: 5 Projects To Get You Inspired

University of Waterloo Real-time Embedded Systems Group - A Gentle Introduction to Generative Adversarial Networks

### ***Press Coverage***

Our research on "Less Than One"-Shot Learning has been featured on [MIT Technology Review](#), [Digital Trends](#), [The Next Web](#), Scientific American (upcoming), and many other outlets.

### **Community Building and Memberships**

***Peer Reviewer - The Stata Journal***

***Peer Reviewer - 2020 IEEE International Conference on Systems, Man and Cybernetics (SMC 2020)***

***PC Member - The 7th IEEE International Conference on Data Science and Advanced Analytics 2020 (DSAA 2020) Special Session: Data Science for Cyber Physical Systems***

***Member - International Neural Network Society***

***Associate Member (Student) - Statistics Society of Canada***

### **Honors**

***Waterloo AI Institute Graduate Scholarship***

***(Jan. 2020)***

***Statistics and Actuarial Science Chair's Award***

***(Sept. 2020)***

***Ontario Graduate Scholarship***

***(Jan. 2021 - Dec. 2021)***

***SSC Student Travel Award for the 2020 SSC Annual Meeting***

***(June 2020)***

***Ontario Graduate Scholarship***

***(Jan. 2020 - Dec. 2020)***

***Statistics and Actuarial Science Chair's Award***

***(Sept. 2019)***

***Math Senate Graduate Scholarship***

***(Jan. 2019)***

***Statistics and Actuarial Science Chair's Award***

***(May 2018)***

***Faculty of Mathematics Scholarship (recurring)***

***(Sept. 2014 - Aug. 2017)***

***University of Waterloo President's Scholarship***

***(Sept. 2014)***