

Danhua Cai

2109-60 Brian Harrison Way
Toronto, On, M1P5J5, Canada

(+1) 6476855766
danhuacai123@gmail.com
<http://danhuacai.com>

Education Background

University of Toronto

B.S. Computer Science (CO-OP) and Statistics (GPA: 3.86/4.0)

Toronto, ON
Sept. 2016 - Now

Awards

E-Fund Scholarship \$5000	2017-2018
E-Fund Scholarship \$5000	2018-2019
E-Fund Scholarship \$5000	2019-2020
UTSC Dean's List	2017-2020

Skills

Programming Languages: Python, SQL, Java, R, Bash, Javascript, MatLab, C, Latex

Operating Systems: Windows 10, MAC OS X, Linux

Interests

Academic: Causal Inference, Data Science, Machine Learning, Statistics, NLP, Deep Learning, AI

Work Experience

- **DSU UTSC library** Toronto, ON
Software Developer Intern *Sept. 2020- Now*
 - Developed an open-source project named mediaCAT, researched websites and Twitter crawlers, transplanted proper Twitter crawlers such as twint and snsrape
 - Re-designed and standardized codes, used Natural Language Toolkit to delete data such as emoticons and special symbols; used Pandas in Python to write scope parser, deepening my understanding of data structure
- **University of Toronto** Toronto, ON
Teaching Assistant(10 TA contracts in Math, CS, Statistics) *Sept. 2018 - Now*
 - Focused on causal inference, Design and Analysis of Data Structures and Calculus I for Mathematical Sciences
 - Attended relevant essay classes to study causal inference, reading extensively in *Causal Inference: What If*, written by Prof. Miguel A. Hernan and Prof. James M. Robins at Harvard University, deepening my understanding in this field
 - Utilized R Shiny app to visualize data to facilitate student understanding
 - Used Leetcode site to practice Algorithms application in Python
 - Engaged in different research realms, including the introduction to computational methods for solving problems in linear algebra, non-linear equations, approximation and integration; design, analysis, implementation and comparison of efficient data structures for common abstract data types; focus on logical reasoning and fundamental notions
- **DDSB, Ontario Ministry of Education** Toronto, ON
Data Analyst and QA Developer *Sept. 2019- Apr.2019*
 - Researched and developed a new machine-learning procedure called Dedupe to improve data matching and detect false positive patterns within data with advanced ML algorithms such as Bayesian Network Classifier, Conventional Neural Network and RNN

- Conducted Automation Testing and Manual Testing and helped finish AODA (Accessibility for Ontarians with Disabilities Act) Testing with the specific toolbar WAT (Web Accessibility Toolbar)
- Utilized Oracle SQL to establish a summary table that can collect the results of all tables and create new queries to analyze the data-matching status

Projects

- **Data Science Research: COVID-19 Impact on Students** Toronto, ON
May. 2020- June 2020
 - Collected submissions and comments from University of Toronto, University of Toronto Scarborough and The University of Tennessee at Martin subreddit using Python crawlers in Jupyter Notebook
 - Performed data analysis from January to June, displaying a total of 10,000 posts across three universities
 - Sorted data using Word Cloud, collected data from Statistics Canada, including the unemployment rate
 - Referenced the data from impacts of the COVID-19 pandemic on postsecondary students, including financial situations and academic plans
 - Drew conclusions about major concerns in student groups (such as tuition fees and online courses) and about whether government supports will relieve some economic pressure
- **Machine Learning Research: Research on Ideologies of Batch Normalization** Toronto, ON
supervised by Prof. Qiang Sun Apr.2020
 - Worked in a group, conducting a deep discussion on batch normalization and layer normalization
 - Formed a hypothesis that batch normalization could reduce internal co-variate shift (ICS), and reviewed other five papers to gather relevant information, ultimately substantiating a different viewpoint
 - Researched how batch normalization helps optimization by analyzing the disagreement on original points of view and the relations between ICS and improvement of training, BN and ICS as well as BN and the smoothness of loss function
 - Concluded that we could add normalization before each hidden layer to make the network more beneficial
 - Investigated further alternatives on normalization and its pros and cons
- **Restaurant Order App Design (Group Project)** Toronto, ON
Feb. 2018- Apr.2018
 - Utilized GUI to write waiters' and manager's user interface; used Strategy Pattern and Iterator Pattern to make this app more efficient and friendly; made the app more efficient in its timing; added new design pattern to the order systems and drew UML diagram for classes

Primary References

Dr. Linbo Wang Assistant Professor, University of Toronto, linbo.wang@utoronto.ca

Dr. Natalia Breuss Associate Professor, University of Toronto, n.breuss@utoronto.ca

Dr. Qiang Sun Assistant Professor, University of Toronto, qiang.sun@utoronto.ca

Dr. Anna Bretscher Associate Professor, University of Toronto, anna.bretscher@utoronto.ca