

# DARRELL AUCOIN

4119 Esplanade Avenue, Montreal, QC, H2W 1S9

Tel: (514) 802-1234, Email: darrell.aucoin@edu.uwaterloo.ca,

GitHub: DarrellAucoin, Kaggle: DarrellAucoin, Personal: <https://daucoin.getforge.io>

## Relevant Work Experience

**Machine Learning Researcher**, Interdata Laboratories (Datalogue), Montreal, QC. May 2017-Aug 2018

- Researched, trained, and deployed Machine Learning models to help automatically prepare data
- Designed machine learning models that will become a product/service:
  - Segmentation machine learning model that breaks a string into it's components
  - Anomaly detection for text columnar data using word and character embeddings
- Improved existing machine learning models:
  - Carried out experiments on using different character mappings for character-wise Convolutional Neural Network helping to increase classification accuracy of models
  - Cleaned data using SQL, and Apache Spark
- Experimented with various machine learning models:
  - Created models designed to help training classifiers when working with noisy data (internship project)
  - Created a model for classifying whole columns of text data
- Investigated layer activation of Neural Networks, helping Datalogue understand its models better

**Software Adviser**, Statistical Consulting Centre, University of Waterloo, Waterloo, ON. 2015

- Provided software advice on R, SQL, and SPSS for graduate students

## Research Experience

**Undergraduate Research Assistant**, Shoja'eddin Chenouri: University of Waterloo. May-Aug 2015

- Worked with Apache Spark on various data projects

**Undergraduate Research Assistant**, Shoja'eddin Chenouri: University of Waterloo. May-Aug 2014

- Investigated and wrote supplementary material on various Big Data Tools

## Data Science Skills

**Modelling:** Studied various machine learning and statistical models through Masters in Applied Computing and Bachelor of Math in Statistics. I also created various models for Datalogue.

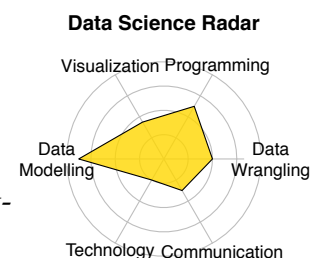
**Data Wrangling:** Cleaned data related to projects using SQL, Spark, R and Python.

**Communication:** Made various presentations to classes, clubs, and conferences.

**Visualization:** Taken a statistics course on Data Visualization.

**Programming:** Experienced in various data analytics programming languages: R, Python, Matlab, and Apache Spark.

**Technology:** Built a small Hadoop cluster using single board computers.



---

**Education**


---

**Master of Science in Applied Computing, Data Science Concentration** 2018

- *University of Toronto, Toronto, Ontario.*

**Relevant Courses:** Topics in ML: Interference & Generative Models, Algorithms for Genome Sequence Analysis, Probabilistic Learning and Reasoning, Fundamentals of Statistical Genetics

**Bachelor of Mathematics with Major in Statistics and Computational Math with CS Minor** 2015

- *University of Waterloo, Waterloo, Ontario.*

**Relevant Courses:** Inference for Big Data, Classification (Machine Learning), Data Visualisation, Function Estimation, Data Types (Python), Object-Orientated Programming (C++), Databases (SQL), Spatial Data Analysis, Longitudinal Data Analysis, Computational Linear Algebra

---

**Projects and Extracurricular Activities**


---

**Automatic Data Cleaning:** Investigated a model to correct noisy labels in training data. 2016

- Probabilistic Learning and Reasoning Course Project

**Imputation of Missing Data using Gaussian Mixtures:** Investigated MCMC gaussian mixture models for data imputation. 2016

- Differentiable Inference and Generative Models course Project

**Review of Genotype Imputation Algorithms:** Compared various genotype imputation algorithms. 2016

- Algorithms for Genome Sequence Analysis course Project

**Statistics Club President, University of Waterloo** 2014-15

- Lead a team in providing various statistics related educational and social activities for fellow students.
- Presented tutorials on SQL, Hadoop, and Big Data (videos up on YouTube Channel "UW Stats Club").

**Small Hadoop Cluster:** Created a small Hadoop cluster using single board computers. May 2015  
Funded by the Math Endowment Fund (MEF).

**Determine Influence in Social Media:** Ranked Twitter users using machine learning algorithms 2014

- Classification Course Project

---

**Honours and Awards**


---

**NSERC Undergraduate Student Research Award:** Natural Sciences and Engineering Research Council of Canada. 2015

- Received 2 awards for work underneath a supervisor at University of Waterloo 2014  
(see Research Experience)

**Winston and Diana Cherry Award in Statistics:** University of Waterloo. 2013

- For highest mark of 98 in a statistics course (Computational Inference)

**Jason Lang Scholarship:** Student Aid Alberta. 2012

- For outstanding academic achievements in undergraduate studies

**Dean's List:** University of Alberta. 2012

**Honourable Mention:** MCM: The Mathematical Contest in Modelling. 2012