

TED HALEY

M.D.S., B.A.Sc.



tedhaley.ca



[linkedin.com/in/Ted-Haley](https://www.linkedin.com/in/Ted-Haley)



github.com/TedHaley



teddyhaley1@gmail.com



604-704-9467

I am a data scientist with experience in research, software development, and engineering consulting. As someone who finds gratification in designing solutions to complex problems, I am interested in finding a position that will challenge me, foster continued education, and promote career development.

EDUCATION

Masters of Data Science (M.D.S.)

University of British Columbia (UBC)

September 2017 – June 2018

- Student Representative
- 4.0/4.3 GPA

Data Analytics

Brainstation

March 2017 – May 2017

Bachelors of Applied Science in Civil Engineering

(Specialization in Management & Transportation Engineering)

University of British Columbia (UBC)

September 2011 – April 2016

- Completed the co-op program
- Member of the Institute of Transportation Engineers
- Member of Engineers Without Borders

KNOWLEDGE & SKILLS

Technical Skills: Packages:

Software Development, Data Analysis, Data Wrangling, Data Visualization, Machine Learning, Statistics, Project Management
Numpy, Pandas, Matplotlib, DASH, Django, Flask, Scikit Learn, Tensorflow, Keras, OpenCV, Airflow, Luigi, PyMongo, SQLAlchemy

Programming Languages:

Python, R, SQL, NoSQL, GIT, BASH, HTML, CSS, Javascript, Java, C++
Kubernetes, Docker, TravisCI, Pip, Anaconda, Tableau, Postgres, MySQL, MongoDB, Jenkins, Jira, Slack, MS Office

Cloud Platforms:

AWS, Google Cloud, Databricks, Linux

WORK EXPERIENCE

Berlinguette Research - UBC

Data Scientist: September 2018 - Present

At UBC, I have been working on developing the autonomous software capabilities of a materials testing robot that utilizes machine learning to accelerate green material design and discovery ([Project Ada](#)). I am a part of a small team of software engineers that operate and develop new features for the robotics systems alongside a group of chemists who design the experiments. My work and responsibilities include:

- Communicate with chemists and project stakeholders to plan for new features and to further develop the capabilities of the system.
- Design and development of automated data processing pipelines, data storage, and data analysis.
- Developing machine learning models to identify important features, classify data, and make predictions.
- Development of a web dashboard for on-the-fly data analysis and visualization.
- Development of multi-objective optimization algorithms for experiment control (Bayesian Statistics).
- Development of convolutional neural networks to classify images and predict features.
- Creating modular services using Docker for continuous integration and continuous delivery of software.
- Managing and deploying to cloud services (GCP, Linux)

BGC Engineering Ltd. (BGC)

Data Science Intern: April 2018 – July 2018

I worked at BGC as a Data Science intern as part of my Masters of Data Science program. There, I was tasked with developing an anomaly detection system and a flood forecasting model. The project included:

- Data wrangling, cleaning, and analysis in Python and R.
- Developing machine learning models for time-series and geographical data that can be scaled for use on thousands of gauges in real-time.
- Measure model performance and compare against the existing system.
- Optimize the ML model to improve accuracy.

RJC Engineers (RJC)

Design Engineer: January 2015 - June 2017

I worked for RJC during and after my undergraduate degree from UBC. At RJC, I was responsible for managing construction projects from proposal to completion. My responsibilities included:

- Site investigation and evaluation of existing building structures.
- Preparation of assessment reports, designs, and contracts.
- Modelling predictions for building maintenance, renewal, and capital projects.
- Communicating with building owners and contractors.

PROJECTS

LightSpeed - DeCode Congestion Hackathon 2019

November 2019: [Demo](#), [Github Repository](#), [Web App](#)

LightSpeed is a remote sensing networking tool intended to be used by a municipality to allow people to dynamically interact with street lights. We developed a working prototype that passively detects pedestrians carrying wifi enabled devices. Using metadata collected by devices located within range of our wi-fi network, we are able to determine the number of devices located within the range of the node and their relative distance to the node.

Project Ada - Berlinguette Research

September 2018 - Present: [Open Source](#)

Project Ada is an automated robotics platform that uses machine learning to accelerate the rate of material discovery. I joined the Berlinguette Research group to lead the development of the data management capabilities of the platform. I work alongside chemists, hardware engineers, and other software engineers to further develop the software requirements of the project. I have developed a data management platform that automates data ETL, data validation and cleaning, data analysis, and visualization. I have also worked with other software engineers to develop a platform to simulate experimental surfaces to benchmark the effectiveness of the optimization algorithms. The tech stack I used is as follows:

- **Database:** NoSQL (MongoDB)
- **Data Pipeline:** Spotify's Luigi
- **Web App:** Plotly Dash, nginx, gunicorn
- **Data Analysis/Machine Learning:** Scikit Learn (regression, clustering, statistics), OpenCV (Image analysis), Tensorflow (Image classification CNN, Predictions, Bayesian Neural Network)

Data ETL, analysis, and the web services are run in Docker images on a Linux based cloud virtual machine. Compute intensive optimization benchmarking and neural network training is done on the Google Cloud Platform.

SCOOT - Personal Project

July 2018: [Open Source](#)

SCOOT is an interactive real-time transportation mapping web application I built using RStudio's Shiny web framework. The intended purpose is to allow the user to get transportation information for walking, cycling, transit, driving, and vehicle sharing platforms. SCOOT queries the location of the user (with their permission), and then queries that information against the Translink API and Google Maps API to get the best route for the users desired mode.

Flow Anomaly Detection System - BGC Engineering

June 2018: *Closed Source*

I developed a real-time anomaly detection system for river flow gauges across North America for BGC Engineering. The system is implemented using a Robust Mean LOESS method in R and Python. This system was able to classify anomalies with an F1 score of 83% on a dataset of over 33 million observations.

2-Day Flood Forecasting System - BGC Engineering

June 2018: *Closed Source*

I developed a 2-day flood forecasting system using Python and Sklearn for BGC Engineering. This system uses historic flow, precipitation, and spatial data to train a Random Forest model. This model was able to classify flood events with an F1 score of 74% on a dataset of over 33 million observations.

Pythia / Karl - UBC

April 2018: [Open Source](#)

Pythia and Karl are Python and R packages that take a dataset with a continuous response variable and various continuous explanatory variables, and provide the user with several functions to build a linear regression model. The fit of the regression can be tested using the diagnosis (residual) plots returned by the plotting function.