Greg **Arbour**

DATA SCIENTIST · STATISTICIAN

Education

MSc. Statistics   *Aug 2020. GPA 4.0*

YORK UNIVERSITY   *Toronto, Canada*

B. Eng. Industrial Engineering   *Apr 2011. GPA 3.2*

RYERSON UNIVERSITY   *Toronto, Canada*

Technical Skills

Programming Proficiencies

- R (advanced), Python (intermediate)
- SQL (basic), Git (basic)

Statistical and Machine Learning Proficiencies

- Proficient in statistical modeling and hypothesis testing, including time series analysis, Bayesian analysis, mixed models and generalized linear models.
- Proficient in contemporary machine learning algorithms, including XGBoost (advanced), deep learning (intermediate), support vector machines (basic), natural language processing (basic).
- Familiar with resampling methods, experimental design and missing data imputation methods.

Experience

St. Michael's Hospital   *Nov 2015 - Oct 2019*

DATA SCIENTIST/QUALITY IMPROVEMENT SPECIALIST   *Toronto, Canada*

- Analyzed large data sets to identify improvement opportunities and then lead the resulting improvement work.
- Modeled length of stay of mental health population using cluster analysis, generalized linear models, time series analysis and propensity score matching.
- Developed predictive model for hospital admission length of stay using free-text doctor notes (using NLP/Topic Modeling, Support Vector Machines & Deep Learning).

Community Care Access Centre   *Feb 2015 - Oct 2015*

QUALITY IMPROVEMENT SPECIALIST (CONTRACT)   *Toronto, Canada*

- Provided coaching and project management to 10+ working groups to execute improvement projects.

Fiix Software   *Jun 2011 - Jan 2015*

PROCESS IMPROVEMENT LEAD/IMPLEMENTATION CONSULTANT   *Toronto, Canada*

- Executed on-site consulting services for software implementation and training.
- Lead efforts to streamline technical support & accounts receivable processes.

Personal Projects

- Contributor to an **R Shiny App** that streamlines Time Series Analysis workflow.
- Developed outlier detection model for length of stay of diabetic patients in large, multi-site hospital using ensemble model approach.
- Used simulation to explore distinction between PSIS-LOO vs WAIC in Bayesian analysis using Rstanarm.
- Contributor to **LKS-CHART Advanced Healthcare Analytics Blog**.