Cameron Porteous *Data Scientist*

📍 Toronto, ON, Canada ✉ cjporteo@gmail.com 📞 226-600-8775 in linkedin.com/in/cjporteo

github.com/cjporteo M medium.com/@cjporteo leetcode.com/cjporteo

Languages & Aptitudes

Python *(Pandas, NumPy, scikit-learn, Keras, TensorFlow, Flask, Matplotlib, Plotly)* • **SQL** • **Jupyter** • **MLFlow**

Linux Shell • **Web Scraping** *(BeautifulSoup, Selenium)* • **Tableau** • **Git** • **Google Cloud Platform** • **Docker**

Technical Skills

Regression (Linear, Lasso, Ridge, Elastic Net, SVR) • **Classification** (Logistic, K-NN, SVM, XGBoost)

Clustering (K-Means) • **NLP** • **Deep Learning** (CNN, RNN, LSTM) • **Hyperparameter Tuning**

Model Interpretation (SHAP Analysis, Partial Dependency) • **MLOps**

Professional Experience

03/2022 – present
Toronto, ON

Data Scientist, *Aviva Canada Inc.* 🔗
- Built personal auto loss-cost pricing model using XGBoost and pyGAM to be deployed in the province of Quebec, responsible for pricing over $85M in annual premium
- Developed end-to-end customer LTV (lifetime value) pipeline for Ontario personal property insurance products, projecting LTV for active policies and incoming quotes
- Engineered performant data pipelines in Python using Luigi to manage data preprocessing workflows (merging, imputation, feature engineering, etc.)
- Designed and conducted A/B experiments to determine customer price elasticity across different segments, improving the overall sophistication of price optimization strategy

02/2021 – 03/2022
Toronto, ON

Statistical Analyst, *Aviva Canada Inc.* 🔗
- Developed Python-based ML solutions for use in territorial risk segmentation, quote conversion, and customer churn modelling
- Worked heavily with gradient boosted tree models, including feature engineering, feature selection, interaction discovery, hyperparameter tuning, and model evaluation
- Produced analytic data visualizations for key stakeholders using tools including Kepler, Plotly, Matplotlib, and Seaborn

07/2019 – 10/2020
Kitchener, ON

Data Analyst, *PWO Canada Inc.* 🔗
- Engineered ETL data pipelines using Pandas to process data for downstream analysis and visualization in Tableau; led to immediate monthly profit increase of $110K
- Automated report generation using Python/SQL data aggregation scripts, vastly improving productivity within the controlling department

Education

2022 – 2024
Atlanta, GA

Georgia Institute of Technology, *Master of Science: Computer Science* 🔗
Specialization in Machine Learning

09/2014 – 04/2019
Waterloo, ON

University of Waterloo, *Bachelor of Mathematics: Statistics; Actuarial Science* 🔗
Double Major; Specialization in Finance

Machine Learning Projects

2020

NBA All-Star Predictor, *(Python, Pandas, XGBoost, SHAP)* 🔗
- Applied gradient boosted tree models to custom web-scraped NBA datasets, classifying NBA All-Stars with an F1 Score of 81.2%
- Leveraged SHAP analysis and partial dependency contour plots to uncover key insights
- Published on Towards Data Science (Medium publication) as a two-part series 🔗

2020

@thoughtfulAi, *(Python, SQLite, GPT-2, TensorFlow, Transfer Learning, GCP)* 🔗
- Generates inspirational quotes using a transformer-based language model (GPT-2) and posts them to Twitter; deployed on GCP Compute Engine 🔗
- Model has been adapted and finetuned to a custom web-scraped dataset