LinkedIn
Github
Email

Aditya Maheshwari

Data Scientist

Profile

Interested in leveraging consumer behaviour to drive product decisions using data science techniques.

Five years of experience across various sectors, working and leading interdisciplinary teams in developing data science applications to improve outcomes.

Education

University of Waterloo / Bachelors of Math, Statistics
Sep 2014 - Dec 2018

Major in Statistics, Minor in Combinatorics and Optimization. Co-operative Program.

Carleton University / Masters of Computer Science
Jan 2019 - Aug 2020

Data Science Specialization. Thesis in Deep Reinforcement Learning to reduce traffic congestion in mixed-autonomy highways. In addition to my degree, I received an A+ in Applied Policy Analysis (paper) and Program and Policy Evaluation (paper).

Public Sector Experience

Federal Government of Canada, Privy Council Office / Senior Data Scientist
Mar 2023 - PRESENT

Enhanced data-driven issue prioritization at the Prime Minister's Office by leading a team of three data scientists to develop an end-to-end machine learning pipeline which extracts topics of interest and sentiments on the PM's emails (~10,000 per day) and automatically refreshes a daily dashboard. More details are shared in the following article.

Techniques: Natural Language Processing, Topic Modelling, Sentiment Classification

Tools: Databricks (Azure), Spark, Python, BERTopic, Fine-Tuning Huggingface Models, PowerBI

Canada Revenue Agency / Data Scientist
Mar 2021 - Feb 2023

Collaborated with UX and program officers to prototype a service that automatically files taxes for ~200K Canadians on social assistance programs, enabling them to receive the social benefits they are entitled to. Received commitment in federal budget 2023 to invest in reducing tax filing burden.

Led four data scientists in developing a synthetic data generation pipeline to publish micro-data usable for research and testing methods without revealing sensitive and confidential information.

Techniques: Data mining, record linkage, synthetic minority oversampling, differential privacy

Tools: SQL, R, Azure

Consulting Experience

Advanced Symbolics Inc. / Data Engineering Consultant
Feb 2021 - Aug 2021

Launched a data pipeline that received clients' topics of interest, extracted and clustered thousands of related tweets, and outputted changes in sentiment and engagement over time.

Techniques: Modularizing Code, Data Pipeline Development, Clustering, Text Cleaning

Tools: AWS, Docker, Apache Airflow, SQL, R, Python, Tableau

Central Bank of Canada / Data Science Consultant
Sep 2020 - Feb 2021

Developed a non-parametric inflation forecasting technique to predict inflation rates within 0.5%.

Techniques: Recurrent Neural Networks (LSTMs), decision-tree-based regression, linking macro-econ datasets

Tools: Python (Jupyter Notebooks, Keras, Scikit-Learn), Tensorflow

Canada Air Transport Security Authority / Data Science Consultant
Sep 2019 - Feb 2020

Improved customer experience by optimizing Canadian airport staff allocation, using a machine learning pipeline to predict the number of passengers arriving for screening in real-time (~5% error).

Techniques: Joined live flight schedule and boarding pass scanner data, applied non-linear regression

Tools: SQL, R (xgboost, recursive least squares)

Private Sector Experience

Interset Inc. (now OpenText) / Data Science Co-op
Jan 2018 - April 2018

Developed and productionized a classification model to detect threats from domain generation algorithm viruses with 90% accuracy.

Techniques: Streaming Billions of Log File Records, Classification, PMML model storage to embed in production

Tools: Apache Kafka, Apache NiFi (Feature Engineering), Random Forest

Scotiabank Capital Markets / Analyst Co-op
May 2016 - Aug 2016, Jan 2017 - Apr 2017

Traded securities for Scotiabank and automated the management of daily funding and repurchase agreements to exchange ~100 million dollars per day and save ~2 hours of an employee's workday.

Worked with traders to evaluate, summarize, and present financial trade data using SQL and Tableau to flag risk tolerance breaches and facilitate strategic trading decisions.

Teaching Experience

University of Ottawa / Introduction to Analytics Managing Instructor
Mar 2023 - PRESENT

Managed three instructors to deliver two 40-hour data analytics courses for 33 upper-year undergrads.

Carleton University / Enterprise Machine Learning Lead Instructor
May 2018 - Aug 2021

Managed two instructors to deliver two 40-hr Enterprise ML boot camps for 55 early-career engineers. Developed curriculum on Python programming, Pandas, Dask, Docker, Kubernetes, and microservices.

Delivered ten 40-hour statistical machine learning boot camps to over 150 upper-year undergrads. Developed curriculum covering NNs, dimension reduction, clustering, classification, and regression.

Supported multiple students with landing jobs in top consulting firms, including Accenture and Deloitte.

Research Projects

Machine Learning for Program and Policy Evaluation
May 2023 - PRESENT

Evaluating the effectiveness of education interventions using data from Camfed studies by applying causal forests to measure treatment effects and clustering to uncover underlying population structures. Submitting a methods paper and an applied paper in Feb 2024.

Posture Up Notification System
Jan 2019 - April 2019

Designed an app to strategically nudge users to correct posture before slouching by applying time series forecasting on data streamed from smartphone motion sensors. Presented poster at Carleton Data Day.

Book Chapters and Medium Articles

Contribution to DS textbook chapters on Clustering, Dimension Reduction, and Data Engineering

Medium Article Series on Artificial Intelligence in the Public Service.

Activities

Performed **cello** internationally in Belgium, Amsterdam, India, USA, Canada; Now just play for fun.
Play **sports,** including volleyball, ice hockey, squash, and running recreationally