

# RAJ CHAUHAN

Halifax, NS B3K 1A1

 [raj.b.chauhan@outlook.com](mailto:raj.b.chauhan@outlook.com) |  [LinkedIn](#) |  [GitHub](#) |  [Portfolio](#)

## PROFILE

---

Data Science professional with proven expertise at an enterprise level. Tenacious, accountable, and self-driven with a Master's degree in Data Analytics and a solid Computer Engineering background. An effective team member who *drives innovation* and *adds tangible value* at every step of Data Science projects, delivering *quantitative outcomes* and *actionable insights* served on time and within budget using technical skills augmented by exceptional communication and presentation skills.

## EXPERIENCE

---

### IT Analyst / Developer - Data Scientist

*May 2022 – Present*

#### Canada Revenue Agency, The Government of Canada

- Contributed to setting up a general framework for *model validation and testing*, along with developing a Responsible AI dashboard in Azure Machine Learning Studio, thereby enhancing the *interpretability and explainability* of ML models.
- Maintained *comprehensive documentation* of all major ML projects and activities, enabling knowledge sharing and informed decision-making, ultimately leading to increased efficiency, effectiveness, and productivity.
- Built a *sequence classification* ML model using *Voting Ensembles* to identify the correctness of notices sent out to taxpayers, and tested its robustness via a *REST API* endpoint, identifying overfitting pitfalls, and necessitating metric re-adjustment and model re-iteration for a better ML model for the business success.
- Conducted *exploratory data analysis and visualization* of financial data (over 1M records) with *PySpark* in *Azure Databricks* as part of GST/HST project, providing valuable insights for strategic business decisions.
- Streamlined *data flow* from replication to loading transformed data into dedicated *SQL* pools with *Azure Synapse Analytics* pipelines and notebooks as a part of Cloud Lakehouse project, resulting in increased efficiency and accuracy of data processing and enabling reliable Machine Learning modeling.

### Data Scientist Co-op / Intern

*Sep 2021 – Apr 2022*

#### Canada Revenue Agency, The Government of Canada

- Contributed to the *research* for determining the usability of *Microsoft Azure Machine Learning* and *Databricks* for an end-to-end ML, furthering the organization's vision of cloud-first approach for protected data workload.
- Delivered a *fully interactive dashboard* using *SAS Visual Analytics* and *descriptive statistics* technique, providing valuable insights into CPU, Memory, Disk and Network usage constraints at the Agency.
- Contributed to *Time-Series Analysis and Forecasting* in R, and predicted future capacity bottlenecks, enabling the Agency to proactively address potential constraints.
- Presented an *end-to-end Machine Learning* productionization framework only using *R* to 100+ Data Scientists of the Data Analytics Working Group, impacting the entire existing ML workflow and streamlining the model development process.

# RAJ CHAUHAN

Halifax, NS B3K 1A1

 [raj.b.chauhan@outlook.com](mailto:raj.b.chauhan@outlook.com) |  [LinkedIn](#) |  [GitHub](#) |  [Portfolio](#)

## EDUCATION

---

**Master of Science in Computing and Data Analytics (GPA – 4.19)**

*Jan 2021 – Apr 2022*

Saint Mary's University, Halifax, NS

**Bachelor of Computer Engineering (GPA – 9.47)**

*Aug 2015 – Jun 2019*

Gujarat Technological University, Gujarat, India

## ACADEMIC PROJECTS

---

### Amazon Reviews Sentiment Analysis

- Developed a sentiment classification model using *PySpark* (Zeppelin) and *SVM* algorithm, achieved 90% accuracy in distinguishing between *positive and negative sentiments of Amazon buyers*, and visualized the top 10 positive and negative sentiments as word clouds in *Tableau*, providing valuable insights into customer sentiment and feedback.

### Canada Sales

- Developed a fully interactive *Power BI* dashboard for demonstrating sales of commercial items in Canada, leveraging *SQL* for data transformations, thereby enabling better decision-making regarding key profitable products and regions within Canada.

### You are what you eat

- Utilized *K-Means* clustering to perform *customer segmentation* using Python, profiling customers based on their engagement with major food categories, providing a deeper understanding of the customer base, and facilitating more targeted messaging to enhance customer engagement.

## CERTIFICATION

---

- [Microsoft Certified: Azure Data Fundamentals](#) *Oct 2022*
- [Microsoft Certified: Azure AI Fundamentals](#) *Jul 2022*
- [IBM Certified: Enterprise Design Thinking Practitioner](#) *Feb 2021*
- [Spark for Machine Learning & AI](#) *Mar 2020*