

RAJ CHAUHAN

>> Data Science | Machine Learning

 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 *add tangible value* at every step of Data Science projects, with *quantitative deliverables* and *actionable insights* served in time and budget using technical skills augmented by exceptional communication and presentation skills.

EDUCATION

Master of Science in Computing and Data Analytics (GPA – 4.19) *Jan 2021 – Apr 2022*

Saint Mary's University, Halifax, NS B3H 3C3, Canada

Bachelor of Computer Engineering (GPA – 9.47)

Aug 2015 – Jun 2019

Gujarat Technological University, Gujarat, India

EXPERIENCE

IT Developer - Data Scientist

May 2022 – Present

Canada Revenue Agency, The Government of Canada

- Conducted *exploratory data analysis and visualization* of financial data (over 1M records) with *PySpark* in *Azure Databricks* as part of Quantum 2.0 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, resulting in increased efficiency and accuracy of data processing and enabling reliable machine learning models.
- Tested a *sequence classification* machine learning model's result for robustness in a *REST API* endpoint, identifying overfitting pitfalls, and necessitating metric readjustment and model re-iteration for more reliable and accurate models.
- Configured Role-based Access Control for various deployed Azure resources, including *ADLS Gen2 Storage containers* and Azure workspaces, ensuring data security compliance and access only to authorized personnel.
- 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*.

RAJ CHAUHAN

>> Data Science | Machine Learning

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

EXPERIENCE (continued)

Data Scientist Intern

Sep 2021 – Apr 2022

Canada Revenue Agency, The Government of Canada

- Delivered a *fully interactive dashboard* using *SAS Visual Analytics* and *descriptive statistics* technique, providing valuable insights into CPU, Memory, Disk and Network usage constraints throughout the Agency, leading to more informed and data-driven decision-making.
- Performed and documented Exploratory Data Analysis, contributed to *Time-Series Analysis and Forecasting* in R, and predicted capacity bottlenecks in the near future, enabling the organization to proactively *address potential constraints and maintain optimal performance*.
- Presented an *end-to-end Machine Learning* productionization solution only using *R* to 100+ Data Science professionals, impacting the entire existing workflow and streamlining the machine learning model development process, leading to increased efficiency and productivity.
- Contributed to the research for *predictive analytics and risk assessment & scoring* using *Databricks*, enhancing the organization's ability to make informed and *data-driven decisions* regarding risk management and mitigation.

PROJECTS

Amazon Reviews Sentiment Analysis

May – Aug 2021

- Developed a *PySpark* (Zepplin) and Hadoop-based *SVM* classification model achieving 90% accuracy in distinguishing between *positive and negative sentiments of Amazon buyers*, providing valuable insights into customer sentiment and feedback.
- Created a *word cloud* in *Tableau* depicting the top 10 positive and negative sentiment words, facilitating quick and easy visualization of key customer sentiment trends.

Canada Sales

May – Aug 2021

- Developed a fully interactive *Power BI* dashboard *demonstrating the sales* of commercial items using various KPIs, including product quantity sold, sales, and profit, as well as data transformations in *SQL*, resulting in better decision-making regarding key profitable products and regions in Canada.

Are Patients Stressed?

May – Aug 2021

- Pre-processed medical records of hospital patients using SQL and implemented an *Artificial Neural Network* algorithm using *Tensorflow* and *Keras* in Python to *predict stress levels of patients* in real-time, achieving 80% accurate prediction, enabling healthcare providers to proactively identify and address potential patient stress levels.