

# Dhvani Soni

Toronto, ON | [dhvanisoni135@gmail.com](mailto:dhvanisoni135@gmail.com) | <https://www.linkedin.com/in/dhvanisoni19/> | [Portfolio](#)

## SUMMARY

**Data Scientist/Analyst** with 3 years of experience in data analysis, machine learning model deployment, and data-driven insights. Proficient in Python, SQL, cloud platforms, and analytics tools to optimize decision-making.

## WORK EXPERIENCE

### Jr Data Scientist

Aug 2023 - Jun 2024

Preemptor AI

Toronto, ON

- **Designed and deployed machine learning models**, achieving a **90%** improvement in predictive accuracy by optimizing data cleaning, transformation, and feature engineering.
- Conducted deep-dive analyses on large datasets, **uncovering key insights** that informed strategic product enhancements and business decisions.
- **Collaborated with engineering teams** to design data pipelines and deploy models, aligning technical solutions with business objectives.

### Data Analyst

Mar 2019 - Apr 2021

Fortune International

Gujarat, India

- Performed **in-depth analysis** using **SQL**, identifying trends and providing actionable insights that led to a **20% improvement** in operational efficiency.
- Designed and maintained **Power BI dashboards** and reports, enabling real-time monitoring and enhancing **decision-making** processes.
- **Automated ETL workflows** and reporting processes, reducing manual effort by **25%** and accelerating data accessibility for stakeholders.

## SKILLS

**Programming Languages:** Python (Pandas, NumPy, sci-kit-learn, TensorFlow, PyTorch), R, SQL

**Cloud Services:** AWS, Azure, Google Cloud Platform, Databricks

**Data Engineering & ETL:** Apache Spark, Hadoop, Airflow, Data Warehousing

**Data and Analytics:** Power BI, Tableau, MS Excel, Data Manipulation, Data Modeling, Data Mining

**AI/ML:** Regression, Classification, Clustering, Time Series Analysis, Neural Networks, Natural Language Processing, Predictive Modeling, LLMs

## PROJECTS

### Student Performance Prediction System

Apr 2024

- Created a student performance prediction system with 96% accuracy, leveraging a Random Forest Regressor model. Designed a CI/CD pipeline using Docker and GitHub action and deployed the solution on AWS.

### Automated Data Pipeline and Analytics with Azure

Feb 2024

- Designed ETL pipelines and workflows using Azure Data Lake, Data Factory, and Databricks, integrating with Azure Synapse Analytics for advanced reporting.

### Bank Transaction Fraud Detection

Aug 2023

- Conducted EDA to analyze trends and identify fraudulent patterns in financial transactions. Designed an unsupervised Isolation Forest model to detect outliers in transaction samples, enhancing fraud detection.

### Sales Performance Dashboard

Jun 2023

- Designed a dashboard in Power BI to monitor sales KPIs, enabling a 15% increase in growth opportunity identification. Extracted and transformed data from a MYSQL database using SQL Queries and performed analysis.

### Predicting Online Payment Default

May 2023

- Built a cost-sensitive XGBoost model with 85% accuracy to predict payment defaults. Handled class imbalance using SMOTE and optimized model performance with feature engineering.

## EDUCATION

### Humber College

Jan 2023 - Aug 2023

Post-Graduate Certificate, Artificial Intelligence with Machine Learning | GPA: 4.0

Toronto, ON

### Durham College

May 2022 - Dec 2022

Post-Graduate Certificate, Data Analytics for Business Decision Making | GPA: 3.8

Oshawa, ON

### Gujarat Technological University

Jun 2014 - Jun 2018

Bachelor's Degree in Computer Engineering

Gujarat, India