

# Arturo Rey

Vancouver, British Columbia



672-558-0655



arturorey.h@outlook.com



arturoreyh



AReyH

<https://areyh.github.io/portfolio>

## Summary

Data Scientist with extensive experience in data modeling using machine learning, business intelligence, and managing the overall data flow, data processing and modeling process. Proven ability to develop business solutions by analyzing and interpreting data to create value. Always eager to learn new bleeding edge technologies.

## Experience

### Data Scientist Consultant

April 2024 – June 2024

*Pacific Salmon Foundation*

*Vancouver, Canada*

- Developed an interpretable predictive model to accurately identify tagged salmon species, aiding researchers in their identification efforts.
- Implemented Outmigration models to forecast salmon migration patterns, enabling researchers to better allocate resources and predict migration timings.
- Designed efficient data pipelines using Apache Airflow and PL/Python with cron, enhancing process automation and data workflow management.

### Data Scientist

April 2022 – December 2023

*Horbath Technologies*

*Barranquilla, Colombia*

- Designed and implemented over 200 predictive models for natural gas distribution stations aimed at forecasting daily consumption using classical and advanced forecasting methods, and Recurrent Neural Networks, allowing the company to better allocate human resources.
- Built risk prediction models using interpretable models that assign a risk level to customers, projected to increase project revenue by 20%.
- Increased customer retention and Diamond customer eligibility through RFM analysis and clustering techniques to recommend proper customer segmentations and improve marketing strategies.

### Engineer

December 2019 – April 2022

*Termocandelaria*

*Cartagena, Colombia*

- Developed a predictive maintenance model using machine learning, significantly reducing maintenance costs and minimizing equipment downtime.
- Managed a team responsible for the operation and maintenance of the plant's production equipment, increasing production uptime, boosting productivity and operational reliability.

## Education

### Master of Data Science

Expected June 2024

*University of British Columbia*

*Vancouver, British Columbia*

### Mechanical Engineering

March 2020

*Universidad del Norte*

*Barranquilla, Colombia*

## Technical Skills

**Languages:** Python, SQL, R, MATLAB, Scala, Javascript, Java

**Technologies:** PyTorch, Tensorflow, Keras, scikit-learn, sktime, Spark, Hadoop, Git, PowerBI, Pandas, Polars, Matplotlib, Seaborn, Altair, Plotly, Docker, Azure, AWS, Apache Airflow

**Concepts:** Artificial Intelligence, Machine Learning, Neural Networks, API, Database Normalization, Agile Methodology, Cloud Computing, Data Engineering

## Projects

### End to End Data Pipeline (ETL) using AWS (Data Engineering 2.0 project) | Python, AWS, SQL, Redshift, PowerBI

- Utilized the CMS.gov API to access and extract OpenPayments data. Orchestrated the seamless transfer and storage of the data within an Amazon S3 bucket, leveraging an EC2 instance for efficient execution.
- Leveraged an EC2 instance to execute the transformation process effectively, guaranteeing data integrity and compatibility for seamless integration and loading into the Redshift database.
- Following meticulous data extraction and transformation procedures, orchestrated the loading of the processed data into PowerBI from the Redshift database. This step ensures accessibility and enhanced usability for data analysts, enabling them to derive actionable insights and make informed decisions effectively.

## Certifications

---

- Databricks Certified Machine Learning Associate
- DS4A: Data Science for All 6.0 (Graduated with Honours) & DS4A: Data Engineering 2.0