Ruwindhu Hettige Don

Entry-Level Data Scientist

+1437-987-6289 ruwindhu123@gmail.com Georgian Dr, Barrie https://www.linkedin.com/in/ruwclk

Summary

Data Science Graduate with a strong background in Software Engineering, and experience in Machine Learning Programming and SQL. Possess excellent analytical and problem-solving skills, and the ability to communicate effectively. Keen learner interested in applying data skills and tools to address a wide range of real-world scenarios.

Experience

Student Program Assistant, Big Data Analytics

Georgian College

Oct 2022 - Present

- Performing program administrative duties, planning student-led events, and assisting in improving the student experience of the Big Data Analytics program.
- Planning and coordinating a student-led Spring 2023 Big Data Analytics networking event.
- Planning and executing industry-partnered projects such as "Big Data Analytics BDAT1011 – Data Analytics Project".

SOFTWARE ENGINEER

IFS Sri Lanka

Aug 2019 to Aug 2021

- Developed complex SQL queries for data extraction and conducted performance optimization in a Manufacturing Planning setting.
- Utilized Oracle SQL Database System including stored procedures, functions, and views.
- Developed proof of concept project using Python to extract Production Scrap data to be used in an existing Machine Learning Model.

SOFTWARE ENGINEER TRAINEE

IFS Sri Lanka

Mar 2018 to Oct 2018

- Analyzed client reported support queries, provided alternate solutions where applicable, ideated feasible solutions for bugs and tested fixes. Upheld strong communication with second-line support and customers.
- Worked collaboratively in multiple agile development and support teams consisting of Business Systems analysts, Software Engineers, Architects and Product Owners.

Education

Georgian College

2022 Sep-Present (April 2023)

Graduate Certificate in Big Data Analytics

- Placed on the Dean's List for Semester 1.
- Relevant modules- Business Processes Modelling, Data Manipulation Techniques, Data Collection and Curation, Data Systems Architecture, Business Intelligence.

Georgian College

2021- 2022

Graduate Certificate in Artificial Intelligence, Design and Implementation

- Placed on the Dean's List for Semester 1 and 2.
- Relevant modules- Machine Learning Frameworks, Machine Learning Programming, AI Decision for Making, Reinforcement Learning Programming.

Staffordshire University (APIIT)

Bachelor of Science in Software Engineering

Second Class Second Division (Hons)

Projects

Optimum Predictive Model Identification for Customer Churn in Music Streaming Industry

- Successfully completed Machine Learning Model building for Capstone Project in the Artificial Intelligence Design and Implementation program, at Georgian College.
- This project focused on using both Machine Learning (Azure Machine Learning Studio) and Neural Network models (Pytorch) to predict the customer churn classification for music streaming services which focus on catering to a niche market.
- Prepared large datasets, created statistical analytical models, and performed model training, validation, and testing.

Spotify Daily Charts Analysis ETL Project

- Extracted global daily chart data from Spotify API using the Spotify package.
- Transformed the chart data using Pandas and used Seaborn to visualize the analysis.
- Stored the cleansed dataset onto a Mongo DB (NoSQL) Atlas database, connected the database to a simple JavaScript WebApp which was hosted on Heroku.

Smart Beach in Kincardine: Wave and Beach Analysis

- Extracted data from publicly available buoy data in Lake Huron from NOAA.
- Performed data cleansing techniques on extracted data.
- Compared prediction performance of multiple Regression Models using scikit-learn for wave and beach condition predictions.
- Conducted multi-output Regression Modelling.

GPU Sales - Market Sentiment Prediction using Twitter

- Acquired GPU Sales related tweets using Twitter API.
- Created Data Pipeline using Apache Kafka and Apache Flume for ingestion and Apache HDFS for Data Storage.
- Developed Machine Learning Models using Spark ML.

Skills

- Strong grasp of applying Python for data science (Pandas, NumPy)
- Proven experience in SQL (SQL Server and Oracle), C# and C++
- Azure Machine Learning, Amazon SageMaker, RapidMiner
- Visual Studio Code, PyCharm and Jupyter Notebooks
- Version Control- Git and SVN
- Microsoft PowerBI and Tableau
- Big data frameworks - Spark ML and PySpark
- Apache Hadoop, Hive, Kafka, Flume for Big Data Systems
- JIRA and Microsoft SharePoint
- **Interpersonal and Teamwork Skills** gained by working in multiple agile teams.
- **Leadership Skills** gained by leading diverse projects during Bachelors and Graduate programs.
- **Program Management Skills** gained by working as a Scrum master in agile development projects.

References

Available on request