Aman Arora

Dartmouth, NS B3A 4W7 | amanarora3dec@gmail.com | +1(437)870-6316 |

www.linkedin.com/in/amanar3ra

Profile

Entry-level Data Analyst with a focus on machine learning and data engineering. Eager to apply skills in **AWS Sagemaker, ELK stack, Python, SQL and Excel** to real-world projects. Demonstrated ability to work with large datasets, extract insights, and build basic predictive models. Seeking a challenging role to contribute to a dynamic team and develop expertise in data analysis.

Education

Nova Scotia Community College - NSCC , Graduate Certificate, Full Stack Application Development — Sept 2024 – April 2025

Humber College , Ontario Graduate Certificate, Artificial Intelligence with Machine Leaning — Jan 2024 – Aug 2024

- GPA: 89/100

Indira College of Commerce and Science, Bachelor of Computer Science — 2020 – 2023

- CGPA: 9.56/10.0

Projects

Music Player — Music$_{Player}$

- A music player application was developed using Python, Tkinter, and Pygame, with a MySQL database for storing playlists. It offers core functionalities like play, pause, next, and previous tracks, and is compatible across various devices.
- Tools Used: **SQL**

Feature Engineering

- Proficient in data preprocessing, transformation, and feature creation. Experience in optimizing model performance through effective feature engineering techniques.
- Tools Used: **AWS Sagemaker**

Data Preprocessing and visualisation

- The ELK Stack and Apache Kafka are great tools for studying Australia's weather, addressing the challenges posed by the 5 Vs of Big Data. By leveraging these technologies, organizations can gain valuable insights and make data-driven decisions to improve their operations and mitigate risks.
- Tools Used: **ELK**

Data Insight Extraction

- I established a Hadoop cluster on Google Cloud Platform (GCP). This distributed computing framework allowed for parallel processing of the data, significantly accelerating the analysis.
- Using Hive, a SQL-like query language for Hadoop, I executed queries sequentially via SSH to extract valuable insights from the data. This approach ensured efficient and scalable data processing.
- Tools Used: **Hive and Hadoop**

Analysis/Prediction of Oceanic Events due to earthquake

- Visualizing our earthquake data reveals patterns that can inform urban planning and disaster preparedness. This data-driven approach helps us mitigate earthquake risks and build more resilient communities.

Speech Sentiment Analysis

- Speech sentiment analysis was conducted using the CREMA-D dataset, which contains 7,442 audio clips of actors expressing various emotions.
- Librosa extracted MFCCs, and audio normalization ensured consistent features. This demonstrates the effectiveness of CNNs for this task.
- Tools Used: **CREMA-D dataset**

Skills/Technologies

Languages: C, C++, Java, React.js, JavaScript, HTML5, Python

Skills: Liner Algebra, Data Validation, Data Analysis, Computational Geometry, Data Processing, Statistical Analysis, Matplotlib, Machine Learning, SQL, Excel, Seaborn, Deep Learning, Big Data.

Certifications and Licenses

- Organizational Culture - Linkedin
- Cloud Fundamentals - ExcelR
- Apply prompt engineering with Azure OpenAI Service - Microsoft
- Build natural language solutions with Azure OpenAI Service - Microsoft
- Robotics and IOT with arduino - Mechatron Robotics
- Cloud Computing basics - Techademy
- Get Started with Azure OpenAI Services - Microsoft
- Introduction to Generative AI - Google Cloud Training Online
- Azure AI Fundamentals - Microsoft