Anindita Guha

Computer Science

708 - 2001 Brunswick Street, Halifax, NS, Canada, B3J 3J7, 902-440-9902, aninditaguha9@gmail.com

Education

Dalhousie University
Masters in Applied Computer Science
[Sep 2019 – Dec 2020 **]**

Parul University
Bachelors in Technology (Computer Science Engineering)
[Jul 2015 – May 2019 **]**

Work Experience

Nova Scotia Health Authority
Data Analytics Intern [Jan 2021 – Current **]**

- Implemented a front-end visualisation panel for management of the hospital resource usages based on statistical analysis and ML.
- Maintaining a streaming pipeline in the Azure Databricks project of the applied analytics department and learning about migration to Kubernetes.

Technologies currently working with: *Python, Azure, Spark, Machine Learning, Tableau*

Dalhousie University
Graduate Teaching Assistant [Sep 2020 – Dec 2020 **]**
Data Management Warehousing and Analytics

- Teaching concepts of DBMS, Analytics, ETL Process, and Visualisation.
- Showcasing hands-on experience with BI Tools, Spark, and R language.

Technologies taught: *Python, R, Spark, Tableau, SQL*

Technical Skills

Data Engineering Skills
Tools: Spark, Hadoop, Tableau, CognosBI
Languages: Python, Octave, R, Java, C++
Database: MongoDB, MySQL, Neo4j
Platforms: AWS, GCP, Databricks
Concepts: ETL, Data warehouse,
Libraries: NumPy, pandas, plotly, matplotlib, Dash, scikit-learn

Other Skills:
Web Development: HTML/CSS, Angular, Node.js, React, Flask, JavaScript, Spring
Concepts: Agile, SDLC, Design Principles, Object-Oriented Programming
Others: GUI Development (Python), Linux, MS Office, Visio, GitHub, CI/CD, JIRA, Heroku

Projects

Spotify Recommendation System [Nov 2020 – Dec 2020 **]**

- Implementation and comparison of different recommendation systems for Spotify playlists for both artists and types of songs.
- Conducted research on existing recommenders, and implemented popularity-based recommenders, and content-based recommenders.
- Have used the KNN algorithm for implementing collaborative filtering and evaluation was performed using precision and recall.

Technologies Used: *Python, matplotlib, pandas, plotly, Scikit-learn*

Global Warming Prediction and Visualisation [Nov 2020 **]**

- A web application that illustrates various visualisation techniques for global climate change occurrences and predictions.
- Used plotly and dash module of python to create visualisations concerning factors of global warming.
- Implemented polynomial regression and linear regression to predict the change in sea level rise/glacier melt with the change in the temperature value and various other factors such as deforestation.

Technologies Used: *Python, Plotly, Dash*

Medical prediction using Decision Tree [Sep 2020 **]**

- The ML approach to diagnosing and predicting the sickness of a patient using their past 4 months activity and weather dataset.
- Conducting research, pre-processing, and applying statistical analysis on the collected big data.
- Applied feature extraction, used the Decision Tree model, and predicted probability for upcoming sickness.

Technologies Used: *Python, Pandas, Scikit-learn*

Data Analysis on Twitter, Movie, and News Data [Mar 2020 **]**
- Twitter, movie, and news data are fetched from API, data cleaning is performed, and finally stored in MongoDB installed in AWS EC2.
- Performed a frequency count on the data using PySpark and stored it in JSON format.
- Data is visualized using a word cloud in Tableau.

Technologies Used: *Python, MongoDB, Spark, Tableau, JSON, AWS EC2*

Learning Management System [May 2020 – Sep 2020 **]**
- Developed a multi-cloud serverless application using AWS and GCP services.
- Created chatbot using Amazon Lex and online chat module using GCP Pub/Sub.
- Performed sentiment analysis of user chats using AWS Comprehend.

Technologies Used: *Python, Flask, React, Nodejs, AWS Services (S3, Comprehend, Lex, Lambda, Cognito, RDS), GCP Services (Pub/Sub, Cloud Storage)*

Personal Attributes

Ethics, Critical Thinker, Creative, Good Listener, Disciplined

Licenses and Certifications

- Enrolled in **Data Engineering with GCP specialization** course with Coursera (To be completed by February 2021).
- Completed **Apache Spark Specialist** in LinkedIn Learning.
- Completed the **Machine Learning** online course at Stanford University.
- Achieved 2nd position in academics, Computer Science Department, Parul University (2018-2019).

Languages

English, Hindi, Bengali, Assamese, Telugu, Gujarati

Profile Links

GitHub: https://github.com/AninditaGuha98

LinkedIn: https://www.linkedin.com/in/anindita-guha-367228109/

Website: http://anindita.codingprivacy.com/