Big Data Analytics Project Proposal

CIS 602: Special Topics in CIS, Summer 2024

Vaishnavi Paineni(38)

Project Title: GHG and Methane Emissions Monitoring and Analysis Using AWS Glue and AWS Quicksight

Project Description:

The "Dynamic GHG and Methane Monitoring and Analysis Using AWS Glue and AWS Quicksight" project aims to establish a comprehensive and scalable system for tracking real-time GHG and Methane emissions. By leveraging Amazon S3 for data ingestion, AWS Glue for data processing, and AWS Quicksight for Analysis, the project will enable continuous monitoring and in-depth analysis of GHG and Methane levels. This system is designed to provide immediate insights and critical information to support environmental management and policy-making efforts.

Project Objectives:

1. Data Acquisition: Implement Amazon S3 Data Stream to consistently collect GHG and Methane emissions data from multiple sources such as industrial sensors, satellite data, and public APIs.

2. Data Preprocessing and Enrichment: Use AWS Glue to clean, normalize, and enhance incoming GHG and Methane emissions data with additional contextual information like geographic and temporal details.

3. Instantaneous Data Insights: Design real-time analytics to measure critical GHG and Methane metrics (e.g., concentration levels, emission trends) for specific regions, providing rapid insights into changes in GHG and Methane emissions.

4. Querying the data through Athena: The data stored in the s3 bucket i.e., raw data and processed data is queried using Athena.

5. Structured Data Storage: Store processed GHG and Methane emissions data in Amazon S3, systematically organized by geographic location, and indexed using AWS Glue Data Catalog for easy retrieval and analysis.

6. Scheduled Data Analysis: Set up AWS Glue jobs to conduct regular batch processing, including historical GHG and Methane data analysis to detect long-term trends and anomalies in emissions.

7. Interactive Data Representation: Utilize Amazon QuickSight or similar data visualization tools to create dynamic dashboards and reports that display real-time and historical GHG and Methane emissions data in an easily interpretable format.

8. Performance Monitoring: Amazon CloudWatch continuously monitors the system's performance and health, ensuring reliability and efficiency.

Conclusion:

The "GHG and Methane Emissions Monitoring and Analysis Using AWS Glue and AWS Quicksight" project aims to provide actionable insights and real-time GHG and Methane emissions data for various applications. Leveraging AWS services, this project will enhance decision-making and safety in scenarios influenced by GHG and Methane levels.

References:

- https://ourworldindata.org/co2-and-greenhouse-gas-emissions
- [Towards Data Science: Free and Reliable Weather Data Sources](https://towardsdatascience.com/five-free-and-reliable-weather-data-sources-20b9ea6afac9)
- [OpenWeatherMap API](https://openweathermap.org/api)
- [NOAA Real-time Data](https://www.noaa.gov/education/resource-collections/data/real-time#:~:text=NOAA%20collects%20real%2Dtime%20data,%2C%20citizen%20scientists%2C%20and%20more)
- [Weather.gov API](https://www.weather.gov/documentation/services-web-api)
- [OpenWeatherMap API](https://openweathermap.org/api)