

CSP554 BIG DATA TECHNOLOGIES ASSIGNMENT-1

Bairi Rohith Reddy
A20526972

[1]

- Google's GFT was built to predict CDC reports, but the algorithm was found to be significantly erroneous, predicting more than double the proportion of doctor visits for influenza-like illness than CDC. This is because the algorithm is mainly dependent on correlation between search data and flu cases but this is not accurate.

[2]

- Big data hubris refers to the assumption that big data is a substitute for, rather than a supplement, traditional data collection, and analysis, without considering quality and sources of data.

[3]

- **It is not always about the size of the data**, GFT has heavily dependent on its google search algorithm and the enormous data generated by its search algorithm, though it provides a broader perspective, and sometimes traditional "small data" often provides information that is not contained in the big data. So, using data from various sources along with expert feedback could improve the flu detection algorithm.

[4]

- Algorithm Dynamics are modifications made by the engineers to the algorithm to improve the performance and overall quality of the algorithm which ultimately improves the commercial service and consumer experience. This is done to ensure the algorithm adapts to the ever increasing data that it gets to process.

[5]

- The aspects of algorithmic dynamics that impacted GFT are the lack of GFT's adaptability to the ever changing search behavior and its failure in understanding the uncertainty in the complex real-world data and finally it is highly dependent on google search algorithm for correlation of flu related terms but it was found that google's search algorithm in the first place is not efficient.

AWS ACCOUNT :-

STEP 1:- CREATING A S3 BUCKET (a20526972-csp554).

The screenshot shows the Amazon S3 console interface. On the left is a navigation sidebar with options like Buckets, Access Points, and Storage Lens. The main content area is titled 'Account snapshot' and includes a 'View Storage Lens dashboard' button. Below this, there are metrics for Total storage, Object count, and Average object size, all showing as 'Pending'. A section titled 'Buckets (1)' contains a 'Create bucket' button and a table with one bucket entry.

Name	AWS Region	Access	Creation date
a20526972-csp554	US East (N. Virginia) us-east-1	Objects can be public	January 14, 2023, 14:24:27 (UTC-06:00)

STEP 2:- UPLOADING AN OBJECT TO THE S3 BUCKET TITLED(a20526972-csp554).

This screenshot shows the 'Objects' tab of the bucket 'a20526972-csp554'. It features a toolbar with buttons for 'Copy S3 URI', 'Copy URL', 'Download', 'Open', 'Delete', 'Actions', 'Create folder', and 'Upload'. Below the toolbar is a table listing the objects in the bucket.

Name	Type	Last modified	Size	Storage class
Tom White Hadoop_ The Definitive Guide, 4th Edition, 2015 - PDF Room.pdf	pdf	January 14, 2023, 14:26:49 (UTC-06:00)	9.6 MB	Standard