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]

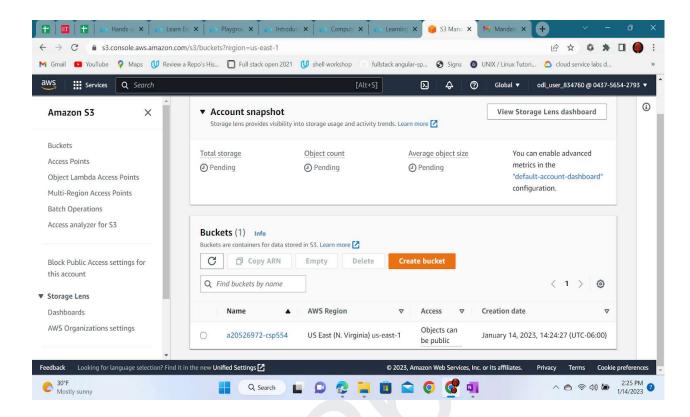
 Algorithm Dynamics are modifications made by the engineers to the algorithm to improve the performance and overall quality of the algorithm which ultimately improves the commercial service and consumer experience. This is done to ensure the algorithm adapts to the ever increasing data that it gets to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



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

