

Rakshith Churchagundi Amarnath

[A20424771]

CSP-554 Assignment 1

Question 3.

Prior Experience:

I have taken courses related to data science like - Data Mining [CS422], Artificial Intelligence [CS480], and Biometrics [CS559]. I have done a summer internship [2019] in BMW Technology Corporation where I worked on Microsoft Azure.

Big Data interests and personal learning goals for the course:

I became fascinated with data when I was working on one of my research paper where I explored the possibilities of detecting pulmonary diseases using Computer vision models by inputting x-rays. My current goal is to learn more about the tools and techniques required to process big data and get a better understanding of it.

Additional topics in the scope of the course of special interest to you:

I am very excited to learn and work on AWS, Kafka, Hadoop and read new articles on the upcoming data processing and visualization techniques.

Other Necessary Accommodations:

None.

Question 5.

Problem with the Google flu detection algorithm – It predicted more than double the proportion of doctor visits for influenza-like illness (ILI) than the Centers for Disease Control and Prevention (CDC), which bases its estimates on surveillance reports from different labs across the US. This happened even though GFT was built to predict CDC reports. The article also explores two issues of GFT – Big Data Hubris and Algorithm dynamics.

Big Data Hubris can be defined as the assumption that **big data** analytics can be used as a substitute rather than a supplement to traditional means of analytics [scientific analysis]. The overestimation of influenza cases by GFT is a prime example for Big Data Hubris.

What approach could have been used to improve the Google flu detection algorithm?

The approach of combining GFT and lagged CDC data, as well as dynamically recalibrating GFT, we can substantially improve on the performance of GFT or the CDC alone.

Algorithm Dynamics – Algorithm dynamics are the changes made by engineers to improve the commercial service and by consumers in using that service.

What aspect of algorithm dynamics impacted the Google flu detection algorithm?

The GFT was an unstable reflection of the prevalence of the flu because of algorithm dynamics affecting Google's search algorithm. One of GFT's error was a media-stoked panic in the last flu season

Question 6

AWS Screenshot:

Bucket Name – a20424771-csp554

The screenshot shows the AWS Management Console interface for S3 Buckets. The left sidebar contains navigation links for Amazon S3, Buckets, Batch operations, Access analyzer for S3, Block public access (account settings), and Feature spotlight. The main content area is titled 'S3 buckets' and includes a search bar, a 'Create bucket' button, and buttons for 'Edit public access settings', 'Empty', and 'Delete'. A summary shows '1 Buckets' and '1 Regions'. Below this is a table listing the buckets.

<input type="checkbox"/>	Bucket name	Access	Region	Date created
<input type="checkbox"/>	a20424771-csp554	Bucket and objects not public	US East (N. Virginia)	Jan 21, 2020 3:55:24 PM GMT-0600

At the bottom, there is a status bar showing 'Operations' with '0 In progress', '1 Success', and '0 Error'. The footer includes a 'Feedback' link, 'English (US)' language selection, and copyright information for Amazon Web Services, Inc.

Folder inside the Bucket - documents

The screenshot displays the AWS S3 console interface. At the top, the navigation bar includes the AWS logo, 'Services', 'Resource Groups', and a user profile 'Rakshith Churchagundi Amarn...'. The breadcrumb trail shows 'Amazon S3' > 'a20424771-csp554'. Below this, the bucket name 'a20424771-csp554' is prominently displayed. A tabbed interface allows switching between 'Overview', 'Properties', 'Permissions', 'Management', and 'Access points', with 'Overview' currently selected. A search bar is present with the placeholder text 'Type a prefix and press Enter to search. Press ESC to clear.' Below the search bar, there are buttons for 'Upload', 'Create folder', 'Download', and 'Actions'. The region 'US East (N. Virginia)' is indicated with a refresh icon. A table lists the contents of the bucket, showing a single folder named 'documents'. The table headers are 'Name', 'Last modified', 'Size', and 'Storage class'. The status bar at the bottom shows 'Operations' with '0 In progress', '1 Success', and '0 Error'. The footer includes 'Feedback', 'English (US)', and copyright information '© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.' along with links to 'Privacy Policy' and 'Terms of Use'.

aws Services Resource Groups

Amazon S3 > a20424771-csp554

a20424771-csp554

Overview Properties Permissions Management Access points

Q Type a prefix and press Enter to search. Press ESC to clear.

Upload Create folder Download Actions

US East (N. Virginia)

Viewing 1 to 1				
<input type="checkbox"/>	Name	Last modified	Size	Storage class
<input type="checkbox"/>	documents	--	--	--

Viewing 1 to 1

Operations 0 In progress 1 Success 0 Error

Feedback English (US)

© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use

Object name – CSP 554 Assignment 1.docx

The screenshot displays the AWS S3 console interface. At the top, the navigation bar includes the AWS logo, 'Services', 'Resource Groups', and a user profile 'Rakshith Churchagundi Amarn...'. The breadcrumb trail shows the path: 'Amazon S3 > a20424771-csp554 > documents'. The bucket name 'a20424771-csp554' is prominently displayed, with an 'Overview' tab selected. Below this, a search bar prompts the user to 'Type a prefix and press Enter to search. Press ESC to clear.' Action buttons for 'Upload', 'Create folder', 'Download', and 'Actions' are visible, along with the region 'US East (N. Virginia)'. A table lists the objects in the bucket:

<input type="checkbox"/>	Name ▾	Last modified ▾	Size ▾	Storage class ▾
<input type="checkbox"/>	CSP 554 Assignment 1.docx	Jan 21, 2020 3:56:45 PM GMT-0600	22.0 KB	Standard

The footer of the console shows 'Operations' with a status of '0 In progress, 1 Success, 0 Error', a 'Feedback' link, 'English (US)' language setting, and copyright information: '© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved. Privacy Policy Terms of Use'.