

Asignatura: OPC13 – Cloud Computing

Ensayo de resultados de aprendizaje de la **semana 9**

Temas: Data Analysis, Big Data, Getting started with compute

Integrantes:

Juan Antonio Díaz Fernandez
Matrícula: 348637
a348637@uach.mx

Diego Alejandro Martínez González
Matrícula: 353198
a353198@uach.mx

1. Resumen Tema “Data Analysis”

El curso presenta el problema con el que cuenta un estudiante llamado Li, su superior del cuerpo estudiantil le ha encargado un problema que consta de averiguar las clases optativas favoritas de los alumnos.

Big Data este concepto se puede describir en términos de desafíos de administración de datos que, debido a un aumento de volumen, velocidad y variedad de datos, no se pueden solucionar con las bases de datos tradicionales.

La profesora que ayuda a Li comenta que es importante identificar los datos que realmente serán útiles para el objetivo que se tenga.

Al momento de ya tener los datos filtrados, Li se pregunta ¿cómo les doy sentido a todos estos datos? La maestra da una opción llamada Kinesis (servicio de AWS que facilita, el procesamiento y análisis de datos de streaming para obtener información oportuna).

Al final se presentan una serie de algoritmos relevantes para este tema y en general para el mundo de la informática:

- Función hash: Acelera las búsquedas diciéndole al programa que pase por alto los registros duplicados.
- Algoritmo de clasificación: Indica cómo agrupar datos.
- Algoritmo de búsqueda: Le indica al programa qué datos recuperar para cada término de búsqueda que un usuario pueda especificar.
- Programación Dinámica: Permite que el programa tome más de una decisión basada en la entrada de un usuario.
- Análisis: Le indica al programa como nombrar a cada pieza de datos.

2. Resumen Tema “Big Data”

El curso comienza explicando que 2 estudiantes están a cargo del baile de primavera en su escuela y quieren usar la información de canciones de sitios web para crear la mejor playlist.

Después el curso enseña el ciclo de vida de gestión de los datos, el cuál es conformado por cuatro puntos:

- **Recopilar y Almacenar:** Implica obtener información precisa de múltiples fuentes y almacenarla.
- **Analizar:** Consta de examinar los datos para descubrir conexiones, sacar conclusiones y tomar decisiones.
- **Visualizar:** Implica representar gráficamente la salida de datos para facilitar la comprensión.
- **Generar:** Se generan más datos con cada transacción o actividad.

Además nos muestra algunos servicios de AWS para esta finalidad:

- **Amazon Machine Learning:** Es una IA que les permite a los sistemas informáticos aprender de los datos.
- **Amazon Rekognition:** Herramienta de análisis de imágenes y vídeos que identifica automáticamente objetos, personas, textos, etc...
- **Amazon Kinesis Firehose:** Se utiliza para capturar, transformar y cargar datos de transmisión en almacenes de datos.

3. Resumen Tema “Getting Started with Compute”

El curso comienza dándonos un concepto general de lo que es el cómputo diciéndonos gracias a este podemos recopilar los datos, analizarlos, procesarlos y distribuirlos, explicándonos también que los recursos principales son el CPU (para ejecutar y procesar las instrucciones), la memoria RAM (para almacenar los datos que se ocupan para llevar a cabo las tareas), el Disco Duro (para almacenar los archivos de arranque y los que se van creando) y el Rendimiento de la red, explicándonos para este último que el ancho de banda es importante para que sea posible enviar una mayor cantidad de datos en un menor periodo de tiempo, al igual nos explica que la latencia debería ser lo menor posible para que se disminuya el tiempo que les lleva viajar a los datos, todo esto es muy importante pues siempre queremos la mayor eficiencia y eficacia posible por lo que con AWS las opciones que otorgan son sumamente útiles para poder llevar a cabo todas las tareas que necesitemos para cumplir nuestros objetivos.