**Propuesta**

**Integrantes**

Dayan Abarca Brenes

Gloria Mata Curling

Nicolas Rojas Rojas

Génesis Villalobos Salas

Marie Sofia Villalobos Martínez

**Descripción**

Desarrollar un sistema que recomiende restaurantes a los usuarios basados en calificaciones, imágenes y opiniones populares de otros usuarios.

**Objetivo general**

El objetivo es crear un sistema que permita a los usuarios descubrir y elegir restaurantes de manera conveniente y personalizada, proporcionando así una experiencia satisfactoria y adaptada a sus preferencias individuales.

**Tipo de base**

Para este proyecto, se utilizará una base de datos NoSQL columnares. Las bases de datos columnares son ideales para el almacenamiento y análisis de grandes cantidades de datos. En este caso, los datos de opiniones de usuarios pueden ser muy grandes, por lo que una base de datos columnares es la mejor opción. Según Morales et al (2019) al organizar los valores en la forma de un vector de atributos permite una fácil compresión de datos y también permite una alta velocidad de escaneo y filtraje. Además, parte de su utilidad es para ser aplicada en tiempo real.

Por otra parte, en el modelado de datos con atributos escasos son ideales, debido a que la mayoría de estos están ausentes en la mayoría de las filas. Esto puede ser común en conjuntos de datos con atributos opcionales o nulos.

La justificación de este tipo de bases de datos es que las bases de datos columnares son más eficientes que las bases de datos relacionales para el análisis de grandes cantidades de datos. Esto se debe a que los datos se almacenan en columnas, lo que facilita el acceso a los datos que se encuentran en una misma columna.

Con respecto a los tipos de datos de la base se espera contar con datos relacionados al restaurante, usuarios que han visitado este, reseñas de estos usuarios, imágenes del lugar y de los alimentos que brinda el servicio, además una calificación promedio de las reseñas que ha tenido.

Seguidamente, se espera que el tamaño de los datos alcance los gigabytes. Esto se debe a que al manejar columnares y una base que incluya información variada de diferentes restaurantes, habría mucha información por lo que la base se volvería pesada, a esto se le llama big data. El big data hace referencia al manejo de muchos datos. Según Morales et al (2019) el mundo digital está creciendo muy rápido y se vuelve más complejo en volumen (terabyte a petabyte), variedad (estructurado, no estructurado e híbrido), velocidad (alta velocidad en crecimiento) y naturaleza. Esto se conoce como el fenómeno global llamado big data. Por ello, el mismo autor propone manejar la alternativa de NoSQL y este volumen de datos.

**Tipo de sistema**

Los sistemas por utilizar son Amazon Redshift y Cassandra. Se seleccionaron debido a que según Ferreira et al (2017) Amazon Redshift es rápido, totalmente gestionable, con una escala petabyte de datos que hace más simple y efectivo analizar todos los datos usando las existentes herramientas de inteligencia de negocios. De igual forma, Cassandra se utiliza bajo las mismas funcionalidades que brinda como Amazon Redshift.

**Bibliografía**

Morales-Morales, M. R., Durán-Cazar, J. W., Tandazo-Gaona, E. J., & Morales Cardoso, S. (2019). Rendimiento de bases de datos columnares. Ingenius, 22, 47–58.

<http://scielo.senescyt.gob.ec/pdf/ing/n22/1390-650X-Ingenius-22-00047.pdf>

Ferreira, P. J., Almeida, A., & Bernardino, J. (2017). Data warehousing in the cloud: Amazon redshift vs Microsoft azure SQL. Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management.

<https://www.scitepress.org/papers/2017/65871/65871.pdf>