

CLASE 39

Visualizar el video:

Principios en la minería de texto

<https://youtu.be/Xab-ZgofNbQ>

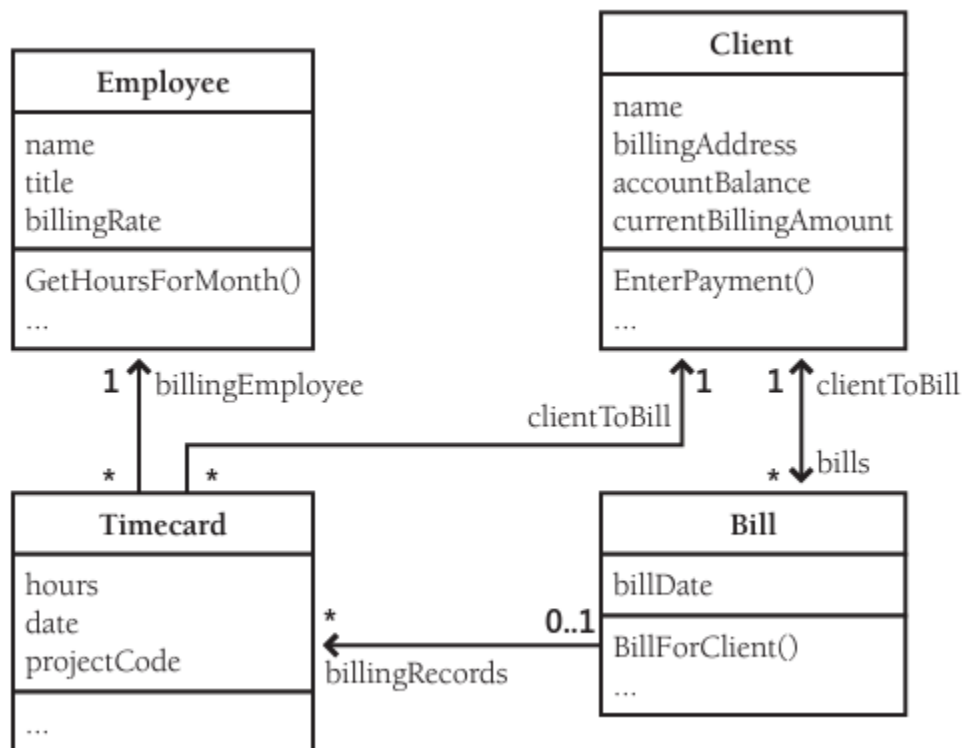
En equipo, apoyados de sus conocimientos en ingeniería de software y con ayuda de la herramienta GIT, desarrollar de manera colaborativa, un sistema distribuido que realiza la búsqueda de frases en un conjunto de textos utilizando tf-idf. El sistema debe contar con una interfaz web como la vista en la práctica 28, o como la vista en la práctica 29 y 30 y con tres servidores de procesamiento adicionales al servidor web.

Para distribuir el trabajo existen dos soluciones principalmente, la primera es distribuir los textos entre los tres servidores, y la segunda es que todos los servidores tengan todos los textos y asignar a cada servidor los textos en los que hará la búsqueda. Se utilizará la segunda opción debido a su escalabilidad.

Los textos donde se realizarán las búsquedas se podrán descargar desde la siguiente liga:

https://drive.google.com/file/d/1BEdeO-yzpTQcbm7_t1am7QKVKdyWYtXD/view?usp=sharing

Para evaluar esta clase se deberá entregar en el chat del equipo en Teams un diagrama de clases del proyecto como el siguiente:



Cuando se termine la aplicación distribuida, se deberá subir al equipo en Teams el código fuente y la liga a la página web para que pueda probar la aplicación y validar su práctica.