

# Algoritmos y Estructuras de Datos

## Trabajo Práctico: Diccionario de Datos

Los diccionarios de datos, también conocidos como tablas de hash, hashtables o hashmaps, son una estructura de datos que asocian claves con un valores para realizar búsquedas de una manera eficiente. Funcionan transformando la clave con una función de hash y obteniendo la posición en la que el valor está guardado.

Las funciones de hash no son a prueba de colisiones, es decir, es posible que haya 2 claves distintas que tengan el mismo hash. En ese caso los 2 o más elementos deben compartir la misma posición en la tabla.

Lo atrayente de estas estructuras de datos es que alcanzan un tiempo constante de búsqueda promedio  $O(1)$  sin importar el número de elementos en la tabla, aunque en casos particularmente malos esta complejidad puede llegar a  $O(n)$ .

## Parte 1

Implementar el TDA `t_diccionario` resolviendo las primitivas antes mencionadas. La implementación del mismo debe ser totalmente genérica, permitiendo que el diccionario almacene datos de cualquier tipo y tamaño, con claves de cualquier tipo de dato.

Se espera que la organización del proyecto esté bien estructurada con los correspondientes archivos de headers (.h) y de código (.c).

Las primitivas del TDA deben ser:

- `crear_dic`: debe recibir por parámetro una capacidad máxima y definir los valores necesarios para inicializar el diccionario.
- `poner_dic`: agrega un nuevo elemento al diccionario. Recibe la clave y el valor a agregar al diccionario, calcula la posición correcta para insertarlo y maneja las colisiones.
- `obtener_dic`: busca un elemento en el diccionario por clave y devuelve el valor.
- `sacar_dic`: Elimina un elemento del diccionario buscando por clave.
- `recorrer_dic`: Recorre el diccionario de datos ejecutando una acción para cada elemento presente.
- `vaciar_dic`: Vacía el diccionario y libera toda la memoria.

## Parte 2

Crear un procesador de textos que a partir de un archivo de texto que ingresa, calcule la cantidad de palabras, espacios y signos de puntuación que contiene, la cantidad de apariciones de una misma palabra, y que genere un podio con las 5 palabras más utilizadas en el texto.

Construya una interfaz de usuario donde se pueda ingresar el archivo a procesar, e informe los datos requeridos de manera clara.

Debe valerse del TDA Diccionario desarrollado en la primera parte.

Debe proporcionar al menos 4 lotes de pruebas, donde se pueda notar cómo funciona el diccionario y donde se vean colisiones en los hashes.

## Entrega y defensa

Este trabajo práctico se realiza en equipos de hasta 4 personas, sin excepción. Los equipos deben ser informados por la plataforma MleL en un mensaje en el foro “Trabajo Práctico” con plazo máximo hasta el día **18/05/2024 a las 23:59**.

La entrega del trabajo práctico se realiza de manera individual. Cada integrante del grupo debe enviar desde la sección prácticas de MleL en la práctica “Trabajo Práctico” a todos los docentes, una copia del programa informando nuevamente en el mensaje los integrantes de su grupo.

El plazo **máximo** de entrega del trabajo práctico será el día **09/06/2025 a las 23:59** a través de la plataforma MleL y habrá una defensa del mismo en la clase de los días 23/06/2025 y 07/07/2025, contando con una evaluación grupal e individual.

Los trabajos entregados después de dicha fecha, no serán tenidos en cuenta, por lo que dicho equipo tendrá un ausente en la entrega y en la defensa, teniendo que utilizar la fecha de recuperatorio para regularizar su situación.