**Sebastián Barrera A00358271**

**Santiago Hurtado A00362570**

**Miguel Sarasti A00364978**

**Sebastián Morales A00365920**

A continuación, se encontrará con una breve introducción al experimento de forma

Iteración 4. Llevar a cabo un experimento que permita comparar los resultados obtenidos por los dos módulos de clasificación desarrollados previamente.

1. **Diseño del experimento.** Definición de cada uno de los elementos propios de un experimento -unidad experimental, factores, variable(s) de salida, repeticiones, niveles y tratamientos-.
2. **Ejecución del experimento.** Aquí debe llevarse a cabo una implementación en C# que al ejecutar en diversas ocasiones los dos módulos de clasificación (muchas -cantidad definida en el diseño-), se obtenga como resultado la(s) variable(s) de salida.
3. **Análisis de los resultados obtenidos.** Lleve a cabo un análisis ANOVA u otro tipo de análisis utilizando herramientas estadísticas.

**Evaluación.** Escriba las conclusiones del experimento teniendo en cuenta los factores estudiados, pero principalmente, la comparación entre la técnica implementada por el equipo de desarrollo y la técnica de la librería.

**Árboles de decisión:** Un árbol de decisión **es un modelo predictivo** que divide el espacio de los predictores agrupando observaciones con valores similares para la variable respuesta o dependiente.

Para dividir el espacio muestral en sub-regiones es preciso aplicar una serie de reglas o decisiones, para que cada sub-región contenga la mayor proporción posible de individuos de una de las poblaciones.

Si una sub-región contiene datos de diferentes clases, se subdivide en regiones más pequeñas hasta fragmentar el espacio en sub-regiones menores que integran datos de la misma clase.

**Desarrollo del experimento**

**Objetivo del experimento**: Determinar la implementación más eficiente entre un árbol de decisión desarrollado por nosotros y el árbol de decisión de una librería (Nombre librería) para predecir la variable objetivo del dataset.

**Unidad Experimental:** La unidad experimental de este experimento son las estructuras de los arboles de decisión tanto el implementado por nosotros y la librería utilizada. Estos serán desarrollados en c# para evaluar su rendimiento, junto a sus pruebas que tendrán factores como el tamaño del dataset que se va a predecir y el acuracy. Con el fin de determinar cuál tiene mejor desempeño.

**Variable(s) de Respuesta:** La variable de respuesta que determinamos para realizar el experimento es el porcentaje de predicción de los datos por cada algoritmo al ejecutarse, cuando se ingrese el dataset.

**Factores controlables:**  Son variables de proceso o características de los materiales experimentales que se pueden fijar en un nivel dado.

* Estado del dataset (en orden aleatorio, ordenado ascendente, ordenado descendente).
* Cantidad de datos de datasetTest
* El tipo de árbol de decisión, tanto el de la librería o el implementado por nosotros.
* La RAM del computador donde se ejecuta el algoritmo (2GB, 4GB, 8GB, 16GB, etc) es un factor que podemos controlar porque podemos escoger computadores con un tamaño de RAM específico o incluso limitar el uso de esta, para el proceso del algoritmo.
* Procesador del computador donde se ejecuta el algoritmo.
* Lenguaje de programación (Java, C#, Python, etc.).
* Sistema Operativo.
* Cantidad de aplicaciones que se están ejecutando en el computador mientras se ejecuta el algoritmo.

**Factores no controlables:** Son factores que no podemos controlar en el experimento.

* Limitaciones del algoritmo a la hora de ingresar cantidades extremadamente grandes datos.
* Cantidad de procesos que se están ejecutando en el computador mientras se ejecuta el algoritmo.
* La cantidad de RAM que utiliza el algoritmo en su ejecución.

**Factores estudiados:**

Los factores que vamos a estudiar en este experimento son los que influyen directamente en la variación del tiempo al momento de llevar a cabo el algoritmo.

* Orden de los elementos por arreglo
* Cantidad de valores por arreglo
* Los algoritmos para estudiar

# Niveles

Tipo de ordenamiento

* Burbuja
* ● Insertion
* Tamaño del arreglo: ● Pequeño
* Mediano
* Grande

Estado de los valores en el arreglo:

* Aleatorio
* Ordenado ascendente
* Ordenado descedente