**Planeación y realización:**

1. **Delimitación del problema:**

Se debe realizar un experimento sobre dos módulos de clasificación, uno de ellos hecho por implementación propia y el segundo bajo una librería. De acuerdo a estas dos implementaciones, se tiene en cuenta el factor de precisión bajo una probabilidad y así identificar que implementación es más viable.

1. **Variables de respuesta y verificación de una medición confiable:**

Las variables de respuesta son la precisión de clasificación de acuerdo al módulo implementado. Y, con el fin de que la medición sea confiable, se repiten las pruebas en diferentes computadores

1. **Factores de estudio:**

Se escogieron como factores de estudio los dos módulos de clasificación, el tamaño del dataset y la RAM del computador donde se ejecuta el algoritmo.

1. **Niveles de cada factor y diseños experimentales:**

En este problema se tienen tres factores de estudio. El primero es el módulo de clasificación, el cual consiste de dos niveles implementación propia y librería ML), el segundo es el tamaño del dataset. El tercer factor de estudio consiste de tres niveles (8GB, 12GB Y 16GB) y es la RAM del computador donde se ejecuta el programa.

1. **Planeación y organización del diseño experimental:**
2. **Realización del experimento:**

**Etapa de análisis:** De acuerdo a la probabilidad dada por cada implementación de los módulos de clasificación, hemos decidido realizar el experimento para determinar qué implementación es más efectiva a la hora de clasificar una película desde el dataset original.

**Etapa de interpretación:**

* De acuerdo al tiempo promedio de entrenamiento de cada uno de los módulos de clasificación, pudimos analizar que la implementación con la librería ML de Microsoft es bastante acertada luego de diversas pruebas comparando con nuestro dataset.
* De acuerdo a la precisión de la selección de películas dados el año, un actor y el género, pudimos analizar que la implementación propia del árbol es bastante acertada luego de diversas pruebas comparando con nuestro dataset.
* Al realizar el análisis ANOVA proponemos una hipótesis nula la cual dice que la implementación propia del árbol y la implementación usando la librería ML no tienen diferencias significativas. Esta, después de una serie de operaciones, es rechazada. Lo que quiere decir que hay diferencias significativas entre las dos implementaciones.

**Etapa de control y conclusiones:**

* De acuerdo a los datos obtenidos, el árbol con mayor asertividad es el de implementación propia.
* Podemos notar que la memoria RAM del computador no afecta el resultado del experimento.
* El árbol realizado usando la librería ML es aproximadamente 20% menos acertado que el de implementación propia.
* El árbol realizado usando la librería ML tiene diferencias significativas al árbol realizado por implementación propia.