НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

Кафедра автоматизації проектування енергетичних процесів і систем

Лабораторна робота № 6

З дисципліни: «Big Data та аналіз даних»

**Виконав:**

студент 5 курсу

групи ТР-02мп, ТЕФ

Круглий Д.В.

**Перевірив:**

Матичин І.І.

Київ – 2021

**Постановка задачі**

1. Імпортувати дані з файлу titanic\_data.csv
2. Усунути змінні home.dest, cabin, name, X та ticket 3. Перетворити змінні pclass і survived на тип фактор.
3. Усунути неіснуючі значення NA.
4. Розбити дані на навчальну та тестову вибірки.
5. Побудувати дерево рішень.
6. Оцінити точність побудованої моделі.

**Хід роботи**

1. Зчитуємо дані з файлу *tatanic\_data.csv* до змінної ***tatanic***.



1. Усуваємо змінні home.dest, cabin, name, X та ticket.



1. Перетворюємо змінні pclass і survived на тип фактор.



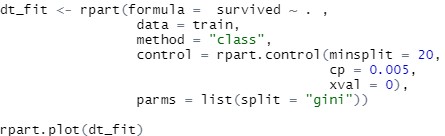
1. Усуваємо неіснуючі значення NA.



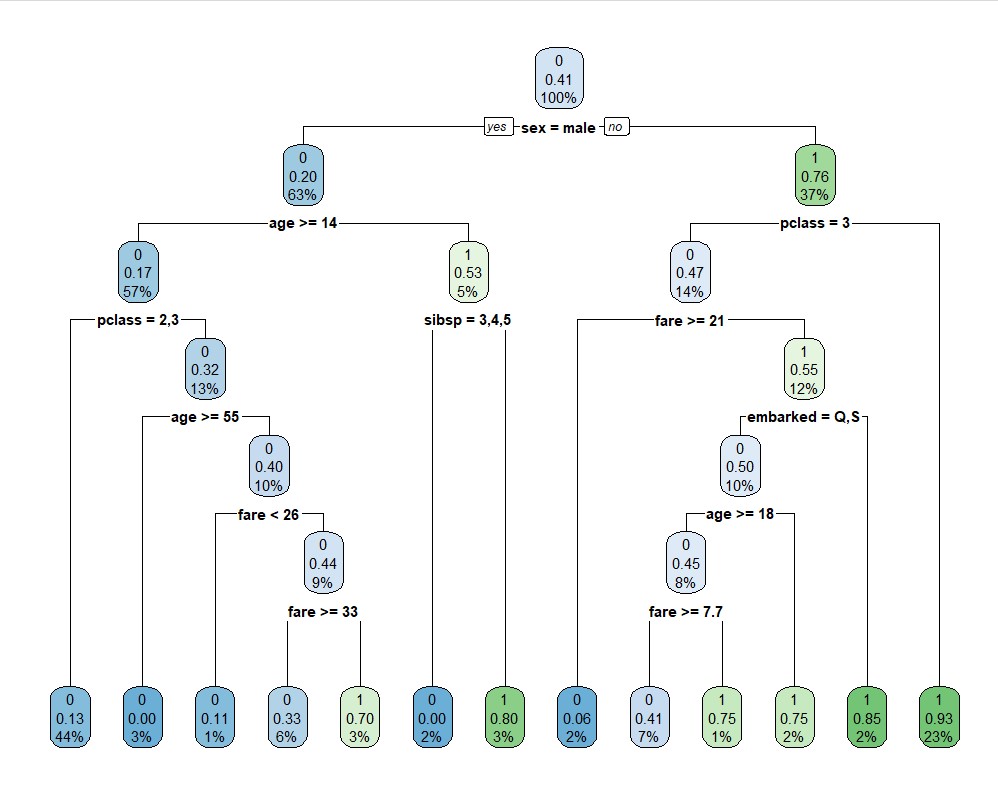
1. Розбиваємо дані на навчальну та тестову вибірки.



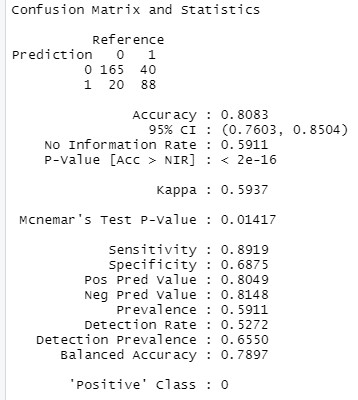
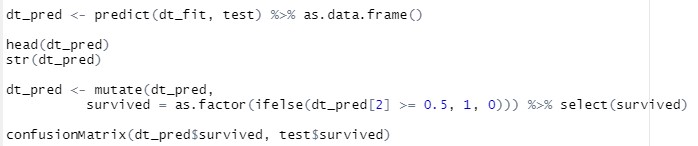
1. Будуємо дерево рішень.



І представляємо отримане дерево візуально.



1. Оцінюємо точність побудованих моделей.



**Висновок:** виконавши дану лабораторну роботу, я отримав навички побудови моделей дерев, їх графічного відображення і оцінювання точності.