**SPRAWOZDANIE**

Zajęcia: Analiza procesów uczenia

Prowadzący: prof. dr hab. inż. Vasyl Martsenyuk

|  |  |
| --- | --- |
| Laboratorium Nr 5  Data 28.04.2023 Temat: Modelowanie procesów uczenia maszynowego w pakecie mlr. Trenowanie, ocena i porównywanie modeli w pakecie mlr Wariant 4 | Maksymilian Grygiel  Informatyka  II stopień, stacjonarne,  Semestr I, gr.1a |

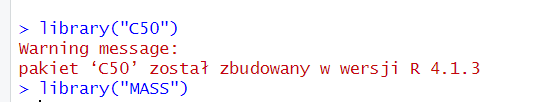
Link do repozytorium: <https://github.com/Maksiolo20/APU>

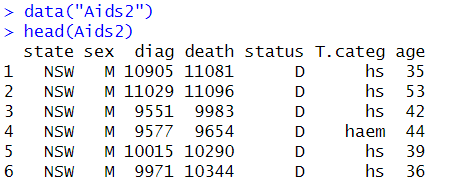
**Zadania:**

**Zadanie 1:**  
Zadanie dotyczy konstruowania drzew decyzyjnych oraz reguł klasyfikacyjnych na podstawie zbioru danych (library(MASS lub datasets)). Wariant zadania nr 4: Aids2

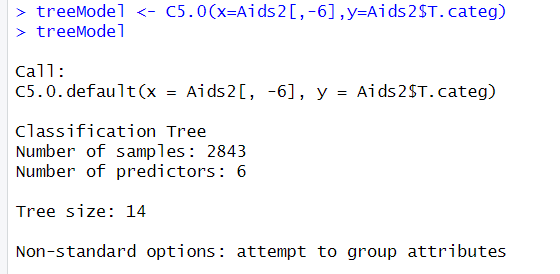
Wykonanie zadania:

Dołączenie pakietów c5.0 oraz Mass:

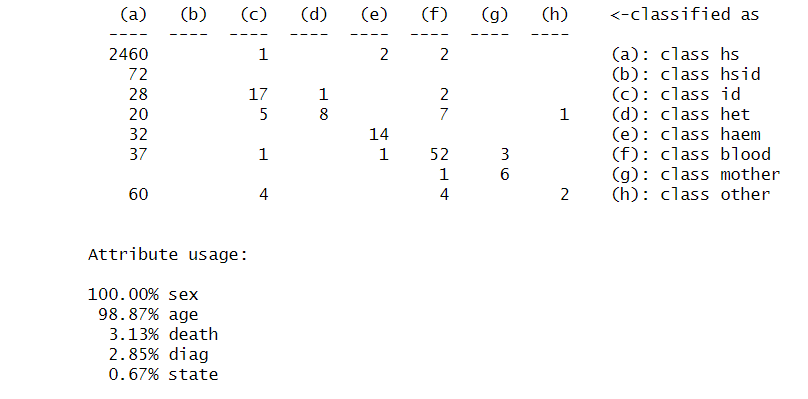
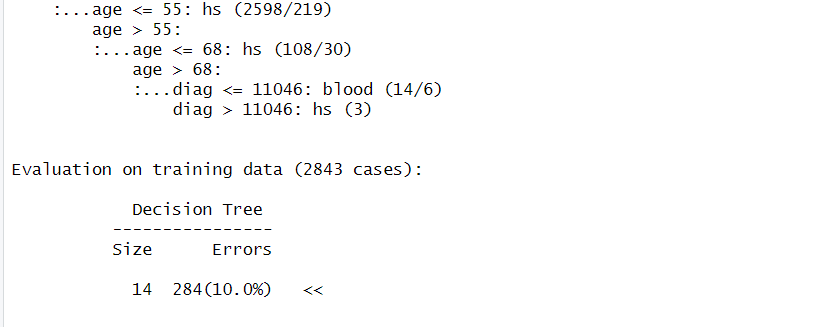
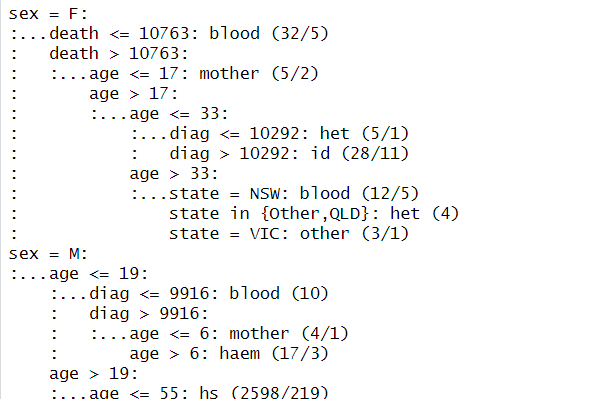
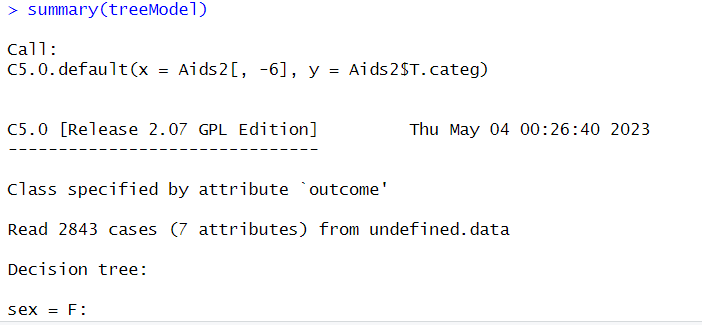


Załadowanie danych oraz wypisanie początkowych danych:

Stworzenie drzewa decyzyjnego I wypisanie go do konsoli:

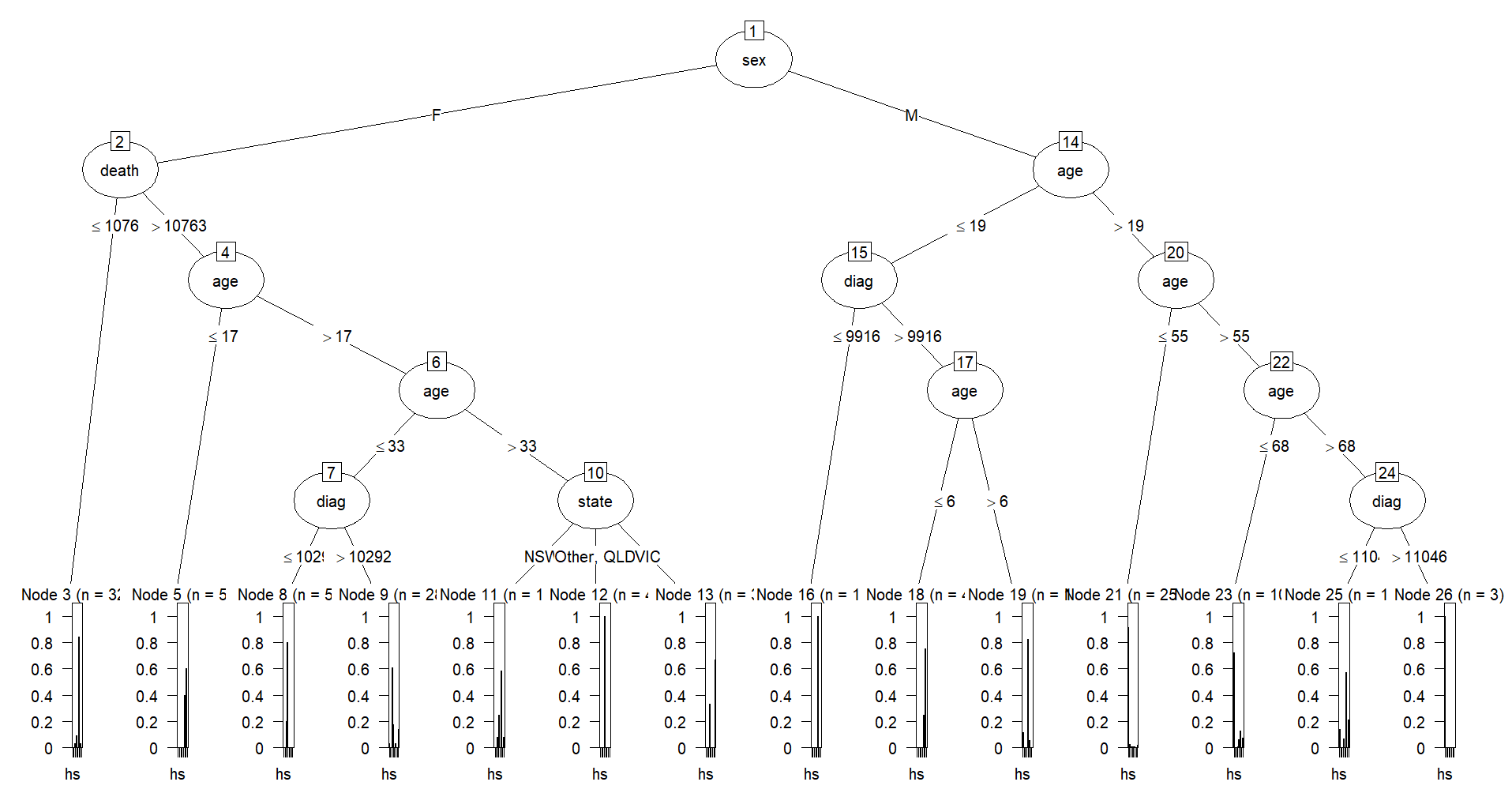


Wypisanie podsumowania uczenia:



Ostatnim krokiem jest narysowanie drzewa w formie wykresu:

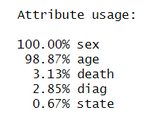




**Analiza wyniku:**

Na podstawie zbioru Aids2 stworzono model drzewa decyzyjnego, które pozwala przewidywać, w jaki sposób dana osoba zaraziła się chorobą Aids. Zmienną która ma być predykowana jest kolumna T.categ - sposób zarażenia.

Dane zostały podzielnoe na kategorie: płeć, wiek, data śmiercim data diagnozy oraz organizacją, do której należy pacjent



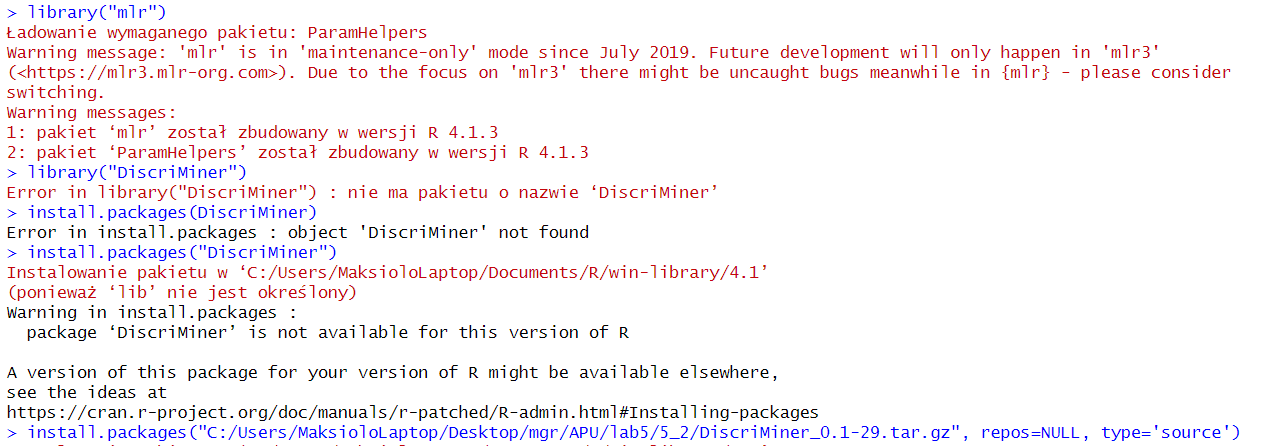
**Wnioski zadania 1:**

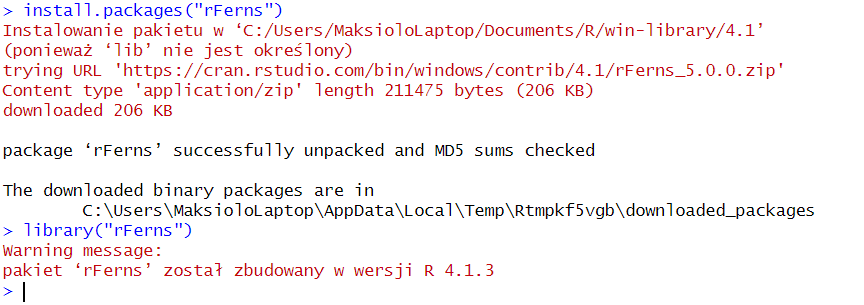
Na podstawie przeprowadzonej analizy można stwierdzić najistotniesjszy element do przewidzenia sposobu zarażenia - płeć. Dla mężczyzn drugin czynnikiem jest wiek, dla kobiet natomiast data śmierci.

**Zadanie 2:**

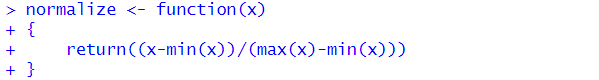
Zadanie dotyczy prognozowania oceny klientów (w skali 5-punktowej, Error < 5%) urządzeń RTV AGD, określonych na Zajęciu 1. Rozwiązanie polega na użyciu pakietu mlr. Należy wybrać najlepszą metodę wśród 5 możliwych z punktu widzenia przecyzyjności. Wyniki porównywania precyzyjności metod należy przedstawić w postaci graficznej.

**Wykonanie zadania:**

Dołączenie pakietów (pakiet DiscriMiner musiał zostać zainstalowany lokalnie):

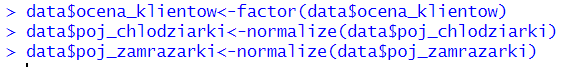


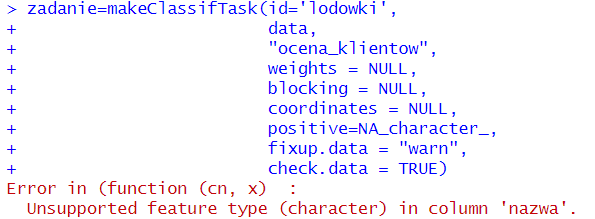
Napisanie własnej funkcji normalizującej:



Zaczytanie danych lodówek:



Wybranie parametrów pod normalizację:  


Na poniższym kroku napotkano błąd “niewspierany typ w kolumnie”

Nie udało się kontynuować zadania, pomimo prób naprawienia błędu.

Zapisanie widoku:

**Wnioski zadania 2:**

Pomimo nieudanej próby rozwiązania zadania, na zajęciach przybliżono działanie prognozowania według określonych parametrów. W przypadku zadania 2 miała to być ocena klientów. Ostatecznie zadanie miało zilustrować 5 metod możliwych I umiejscowić je odpowiednio na diagramie, w zależności od precyzyjności. Przykładowe rozwiązanie w formie graficznej:  
