

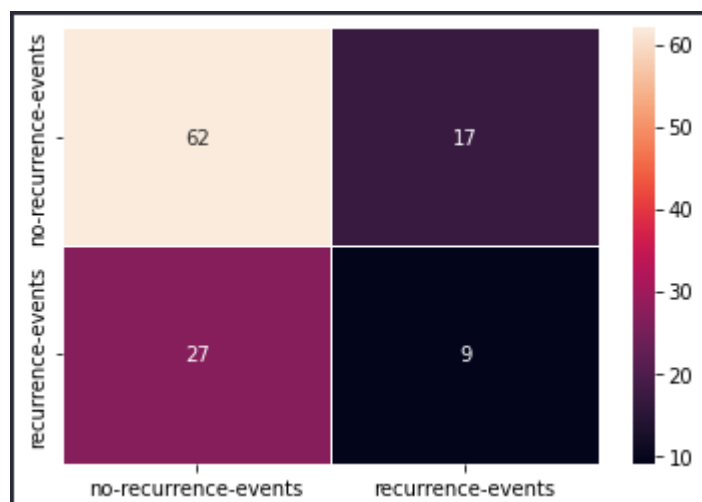
Wprowadzenie do sztucznej inteligencji - ćwiczenie 4

Zadanie

Zaimplementować klasyfikator ID3 (drzewo decyzyjne). Atrybuty nominalne, testy tożsamościowe. Podać dokładność i macierz pomyłek na zbiorach: Breast cancer i mushroom. Sformułować i spisać w kilku zdaniach wnioski.

1. Zbiór breast-cancer

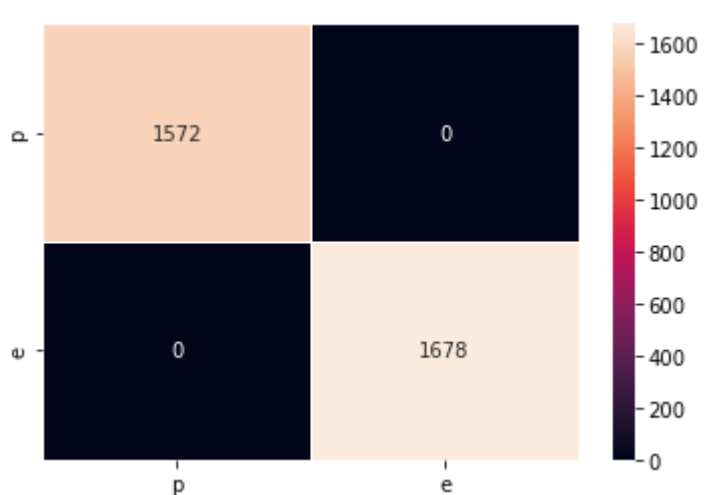
Accuracy: 0.6173913043478261



(in oY - real values, in oX - predicted values)

2. Zbiór agaricus-lepiota

Accuracy: 1.0



(in oY - real values, in oX - predicted values)

Wnioski

Możemy zauważyć, że zbiór numer 2 osiągnął znacznie lepsze rezultaty. Wiąże się to zapewne ze znacznie większą liczbą danych jaka zastała wykorzystana do trenowania drzewa. W zbiorze numer 1 znacznie

częściej wskazywana była wartość 'non-recurrence-events'. Jest to spowodowane tym, że w zbiorze testowym znajdowały się gałęzie, które prawdopodobnie nie zostały wykorzystane do trenowania drzewa. W takich przypadkach zwracana była najczęściej występująca wartość w danym poddrzewie.