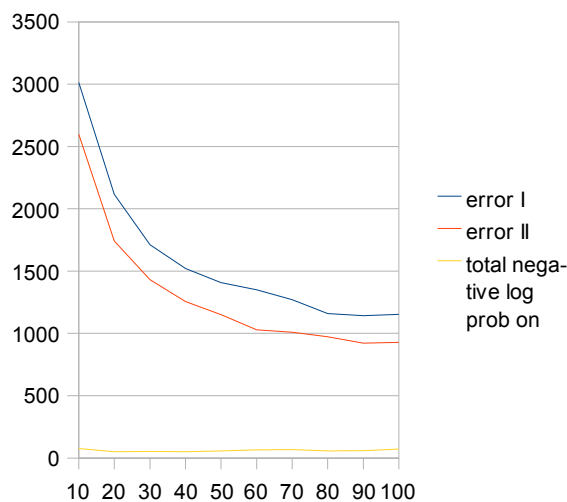


## Pytanie 1.

Zmieniając liczbę jednostek ukrytych na wartości z przedziału  $\{1,5,10,20,50,55,60,75,100,200\}$ , przy liczbie iteracji równej 100, możemy zauważyć, że wraz ze wzrostem liczby jednostek ukrytych, suma kwadratów różnic pomiędzy danymi i danymi odtworzonymi przez jednostki ukryte dla obydwu modeli sukcesywnie maleje, natomiast przy ilości jednostek ukrytych równej 55 dane najmniej zachodzą na siebie (czyli wartość total negative log prob on test data jest najmniejsza i jest równa 55.07 bitów), co oznacza najlepszą separację obydwu klas.

Sprawdzając wpływ ilości iteracji treningowych, testowałem model przy wykorzystaniu ilości jednostek ukrytych równej 55. Najlepszą separację obu klas otrzymałem dla ilości iteracji treningowych równej 40, wartość total negative log prob on test data jest równa wtedy 49.89 bitów. Zraz ze wzrostem ilości iteracji treningowych, maleje suma kwadratów różnic pomiędzy danymi i danymi odtworzonymi przez jednostki treningowe dla obydwu modeli, analogicznie jak przy zmianie liczby jednostek ukrytych.

Wykres zależności pomiędzy ilością iteracji a wartością błędów:

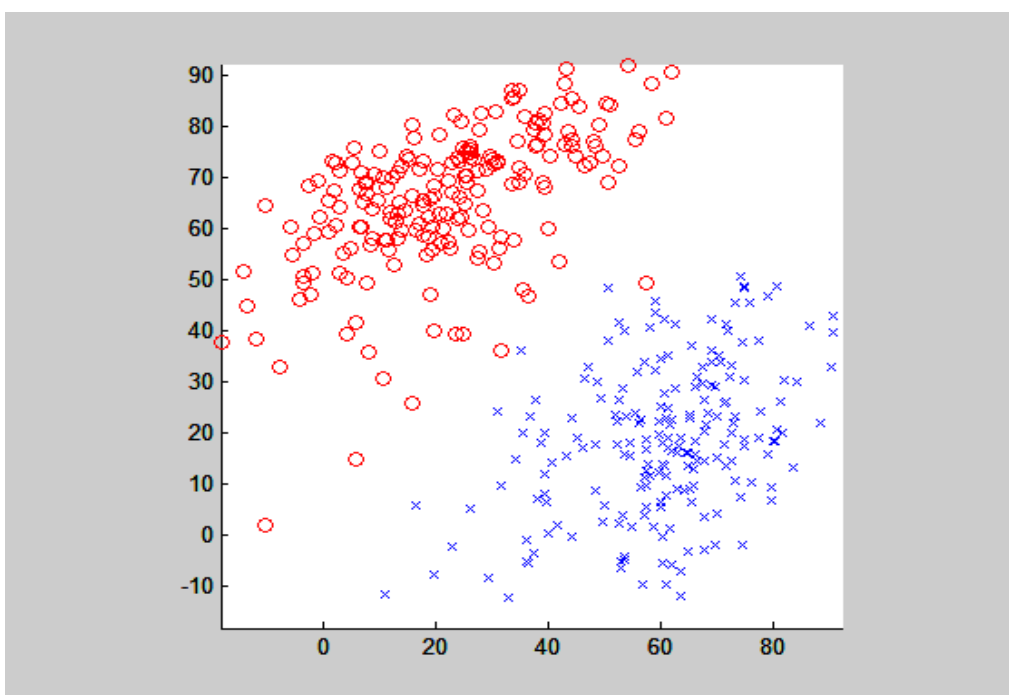
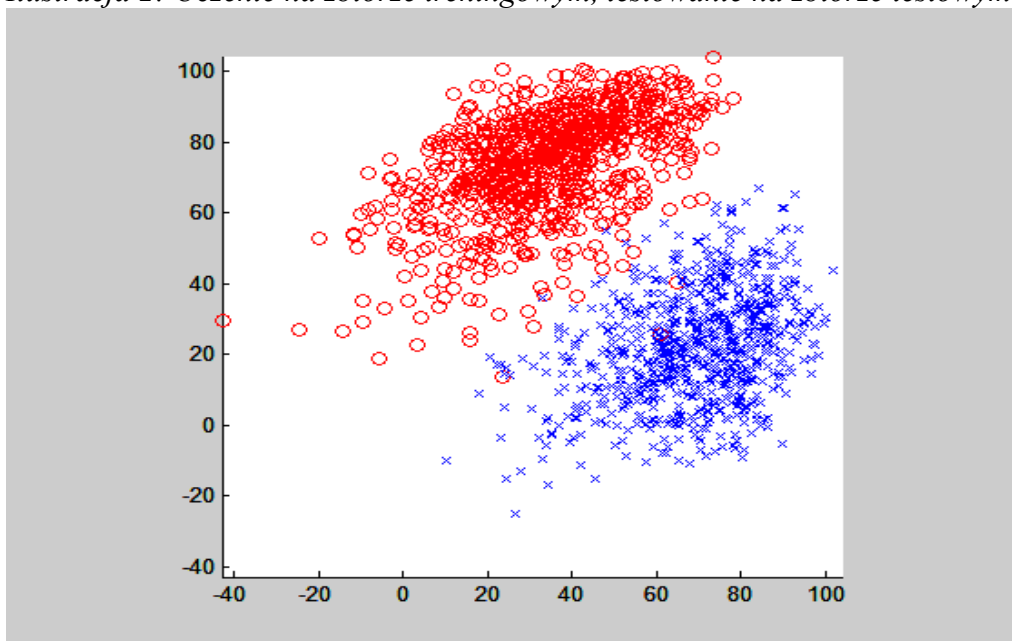


*Ilustracja 1: Zależność wartości błędów od ilości iteracji*

## Pytanie 2.

Ucząc model na zbiorze testowym, jak i treningowym, przy założeniu, że liczba jednostek ukrytych jest równa 55 a liczba iteracji 40, możemy zaobserwować, że zamiana zbiorów prowadzi do lepszej separacji zbiorów, co jest odwzorowane w wartości total negative log prob on test data – dla modelu uczonego na zbiorze danych treningowych jest on równy 49.89 bitów, dla modelu uczonego na zbiorze danych testowych zaledwie 6.73 bitów – jest to spowodowane dużą dysproporcją liczebności zbiorów treningowego (800 przypadków) i testowego (200 przypadków).

*Ilustracja 2: Uczenie na zbiorze treningowym, testowanie na zbiorze testowym*



*Ilustracja 3: Uczenie na zbiorze testowym, testowanie na zbiorze treningowym*