

Metody Inteligencji Obliczeniowej - Sprawozdanie lab 1

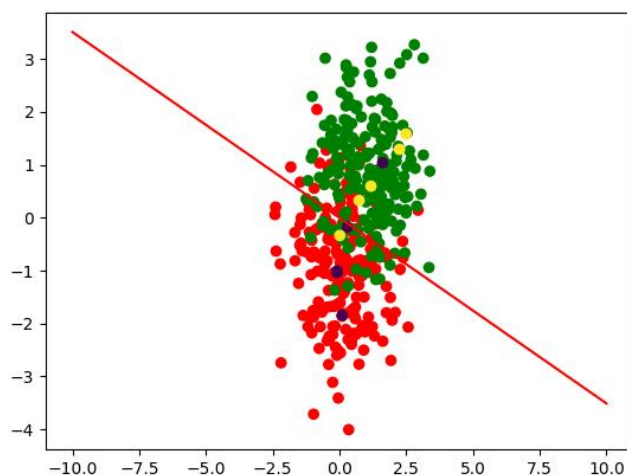
Karol Hoerner de Roithberg

13 Marzec 2023

1 Klasyfikacja

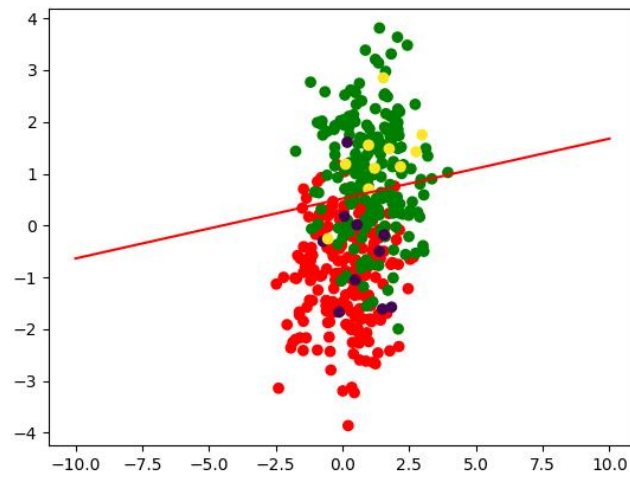
W pierwszym zadaniu należało przeprowadzić klasyfikację dwóch klas pochodzących z rozkładów normalnych $N([0, -1], 1)$ i $N([1, 1], 1)$. Rozkład przeprowadziliśmy dla ilości punktów uczących $n = [5, 10, 20, 100]$, liczba punktów testujących jest równa $i = 200$

Wynik = 0.375.



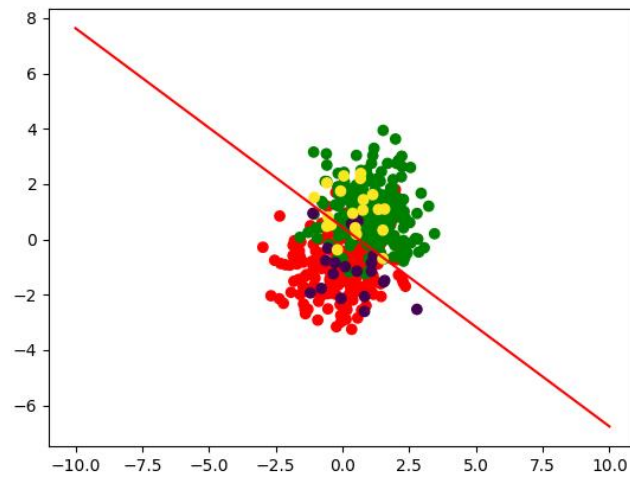
Rysunek 1: Wykres klasyfikacji dla $n = 5$

Wynik = 0.8425.



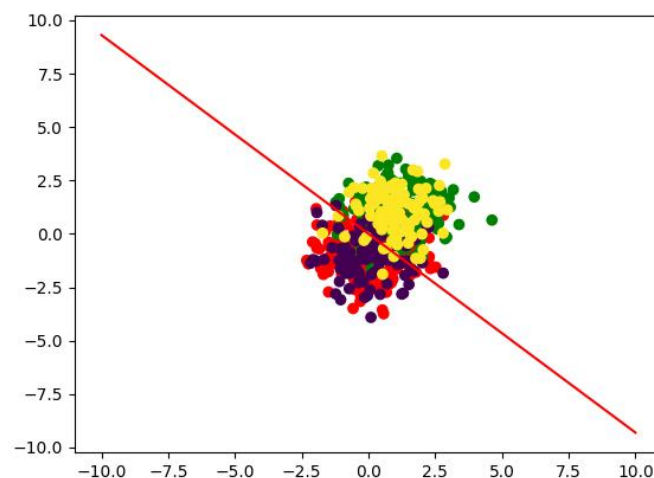
Rysunek 2: Wykres klasyfikacji dla $n = 10$

Wynik = 0.84.



Rysunek 3: Wykres klasyfikacji dla $n = 20$

Wynik = 0.8425.



Rysunek 4: Wykres klasyfikacji dla $n = 100$

1.1 Wnioski

- Skuteczność klasyfikacji rośnie wraz z ilością punktów uczących. Im więcej punktów uczących, tym bardziej hiperpłaszczyzna będzie dopasowana do danych, co przekłada się na lepsze wyniki klasyfikacji.
- W celu klasyfikacji danych na dwie klasy, używamy hiperpłaszczyzny do oddzielenia danych jednej klasy od drugiej.
- Wykres danych i hiperpłaszczyzny jest przydatny do wizualizacji wyników klasyfikacji i pomaga zrozumieć, jak działa algorytm klasyfikacji.

2 Analiza próbek benzyny

W kolejnym zadaniu został użyty algorytm Perceptronu do klasyfikacji próbek czystości benzyny na dwie klasy. Algorytm ten został zastosowany na dwóch modelach Perceptronu, z których jeden działał bez zmian, a drugi miał ograniczoną liczbę iteracji. Celem algorytmu było znalezienie hiperpłaszczyzny, która oddziela dwie klasy na podstawie danych treningowych.

Wynik dla domyślnego parametru: 0.466

Wynik dla 5 iteracji: 0.7

2.1 Wnioski

- Domyślne parametry w tym przypadku dają gorszy wynik niż ograniczenie iteracji nauki Perceptronu, co nie jest przewidywane. Czynnikiem może być kilka, np: zły model danych lub losowość
- Wynik Perceptronu z domyślnymi ustawieniami jest bliski 50%, co wskazuje na złe wytrenowanie modelu

3 Zbiór Irysów

3.1 Macierze pomyłek

W kolejnym zadaniu przeprowadzono analizę macierzy pomyłek dla pięciu różnych podziałów zestawu irysów na dane uczące i testujące.

```
Score: 0.5666666666666667
Coffusion matrix:
[[ 6  2  0]
 [ 0 10  0]
 [ 0 11  1]]
Score: 0.5666666666666667
Coffusion matrix:
[[ 3  5  0]
 [ 0 12  0]
 [ 0  8  2]]
Score: 0.9333333333333333
Coffusion matrix:
[[ 9  0  0]
 [ 0 10  2]
 [ 0  0  9]]
Score: 0.7333333333333333
Coffusion matrix:
[[11  0  0]
 [ 5  0  3]
 [ 0  0 11]]
Score: 0.8333333333333334
Coffusion matrix:
[[10  0  0]
 [ 0 10  0]
 [ 0  5  5]]
```

Rysunek 5: Wynik i macierze pomyłek dla zestawu irysów w 5 iteracjach

3.1.1 Wnioski

- Wykonano wiele losowych podziałów próbek na zbiory treningowy i testowy w celu uzyskania bardziej wiarygodnych wyników. Takie podejście umożliwiło uśrednienie wyników i uniknięcie zależności wyników od jednego konkretnego podziału próbek.
- Dokładność klasyfikacji różniła się w zależności od konkretnego losowego podziału próbek, oscylując w zakresie od 0,57 do 0,94. Takie rozbieżności mogły wynikać z małej liczby próbek w niektórych klasach lub z losowych zbiegów i rozbiegów w procesie uczenia.

3.2 Podział zbioru irysów

W kolejnym zadaniu należało podzielić zbiór irysów na 3 sposoby i zbadać ich wpływ na wyniki.

```
Test/Train: 0.1/0.9  
Score: 0.6666666666666666  
Test/Train: 0.3/0.7  
Score: 0.6666666666666666  
Test/Train: 0.5/0.5  
Score: 0.7733333333333333
```

Rysunek 6: Wyniki dla danych podzielonych na 3 sposoby

3.2.1 Wnioski

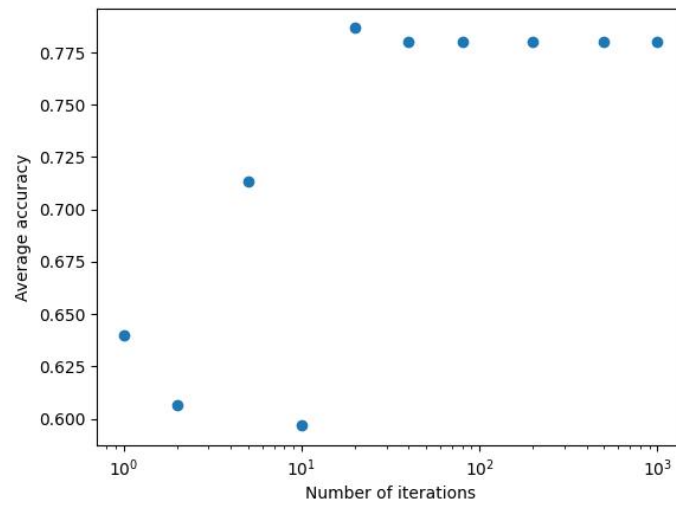
- Dokładność klasyfikacji zależy od ilości danych trenujących, im więcej tym lepszy wynik.
- Aby uniknąć niedokładności należy dobrać odpowiednią ilość danych treningowych.

3.3 Liczba epok

Ponownie posłużono się zbiorem irysów, aby zbadać zależność między ilością epok, a wynikiem.

```
Liczba epok: 1 Wynik: 0.64  
Liczba epok: 2 Wynik: 0.6066666666666667  
Liczba epok: 5 Wynik: 0.7133333333333334  
Liczba epok: 10 Wynik: 0.5966666666666666  
Liczba epok: 20 Wynik: 0.7866666666666667  
Liczba epok: 40 Wynik: 0.7799999999999999  
Liczba epok: 80 Wynik: 0.7799999999999999  
Liczba epok: 200 Wynik: 0.7799999999999999  
Liczba epok: 500 Wynik: 0.7799999999999999  
Liczba epok: 1000 Wynik: 0.7799999999999999
```

Rysunek 7: Wyniki dla Perceptronu o różnej ilości epok



Rysunek 8: Wyniki dla Perceptronu o różnej ilości epok przedstawione na wykresie

3.3.1 Wnioski

- Zwiększenie ilości epok daje dokładniejsze wyniki.
- Przy stosunkowo wysokiej liczbie epok, kolejne ich zwiększanie nie ma dużego wpływu na poprawność klasyfikacji.
- Wykres wyników w zależności od ilości epok przypomina nieco wykres logarytmu.