

Sprawozdanie

Klasyfikacja danych z użyciem neuronu dyskretnego

Amadeusz Filipek

Laboratorium komputerowe WFiIS AGH

1. Wstęp

Celem ćwiczenia jest zastosowanie pojedynczego neuronu typu perceptron do problemu klasyfikacji danych.

Sztuczna sieć neuronowa to rodzaj matematycznego modelu, który ma za zadanie procesować informacje. Sieci neuronowe wchodzą w skład dziedziny nauki określonej jako nauczanie maszynowe (machine learning). Sieć neuronowa zbudowana jest z neuronów ułożonych w kolejne warstwy i połączonych ze sobą. Pojedynczy neuron uzyskuje wejściowe informacje w postaci wektora wejść \vec{x} i procesuje je oddając wartość wyjściową y . Zarówno wartości wejściowe jak i wyjście są liczbami rzeczywistymi. Każda wartość wejściowa przemnażana jest przez wagę przypisaną do danego wejścia. Tak przemnożone wejścia są następnie sumowane:

$$net = \sum_{i=1}^N w_i x_i + b$$

gdzie b – bias stanowi wartość zerowego wejścia o wadze zawsze równej 1. Wartość bias wprowadzana jest aby przesunąć wartość net o pewien offset, co w niektórych przypadkach jest niezbędne do poprawnej pracy neuronu. Wartość wyjściowa y uzyskiwana jest na podstawie funkcji aktywacji neuronu i w najprostszym przypadku ma postać funkcji skokowej Heaviside'a:

$$y(net) = \begin{cases} 1 & net \geq 0 \\ 0 & net < 0 \end{cases}$$

Praca neuronu w głównej mierze zależy od wag. Aby neuron wykonywał powierzone mu zadanie np. klasyfikacji zbioru danych, musi mieć odpowiednio dostosowane wartości wag. Wagi neuronu są dostosowywane w trakcie procedury nauczania sieci. Jedną z metod jest reguła delty, która polega na pokazaniu neuronowi zbioru danych dla którego znana jest odpowiedź. Wagi dostosowywane są na podstawie różnicy pomiędzy aktualnymi wynikami sieci a znanymi, poprawnymi odpowiedziami. Ta procedura nauczania wyróżnia perceptron od sieci neuronowej typu ADALINE, w której wagi uaktualniane są na podstawie zagregowanej wartości net a nie na podstawie wyjść y neuronów.

2. Realizacja ćwiczenia – wygenerowane losowo dane

W ramach ćwiczenia wygenerowano dwa zestawy danych – zestaw uczący oraz zestaw testujący. Dane podzielono na dwie klasy, które oznaczono klasą 0 oraz klasą 1. Każda z klas opisana jest dwoma cechami. Wartości cech dla klasy 0 wygenerowano na podstawie rozkładów normalnych o wartościach oczekiwanych [0, 0] oraz wariancjach [1, 1]. Wartości cech dla klasy 1 wygenerowano na podstawie rozkładów normalnych o wartościach oczekiwanych [3, 0] oraz wariancjach [1, 1]. Klasa 1 różni się zatem cechą pierwszą, której wartość oczekiwana przesunięta jest o 3 w stosunku do klasy 0. Zbadano jak zmienia się błąd klasyfikacji danych za pomocą pojedynczego neuronu typu perceptron w zależności od ilości danych w zestawie uczącym. Uzyskano wyniki dla zestawów uczących po 5, 20, 50 oraz 100 elementów każdej z klas podczas gdy zestaw testujący zawierał po 500 elementów obu klas.

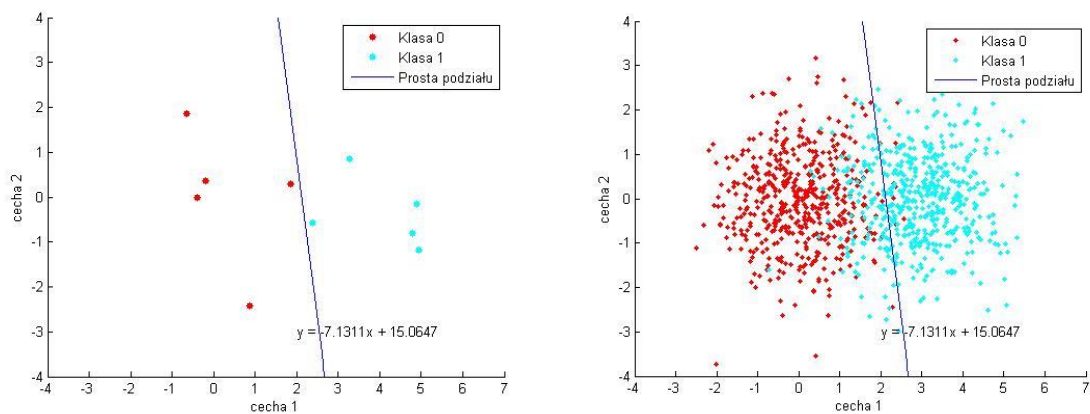
a) 5 elementów obu klas w zbiorze uczącym

Perceptron nauczony zestawem po 5 elementów z obu klas uzyskał następujące wyniki. Poniżej przedstawiono tabelę pomyłek klasyfikacji elementów zarówno dla zbioru uczącego jak i testującego. Neuron pomylił się w 0% dla danych uczących oraz w 11.5% dla danych testujących.

Tabela 1. Macierz pomyłek klasyfikacji perceptronu przy zestawie uczącym po 5 elementów

	Dane uczące zakwalifikowane do klasy 0	Dane uczące zakwalifikowane do klasy 1	Dane testowe zakwalifikowane do klasy 0	Dane testowe zakwalifikowane do klasy 1
Klasa 0	5	0	491	9
Klasa 1	0	5	106	394

Następnie sporządzono wykres przedstawiający położenie elementów klas w przestrzeni cech wraz z linią wyuczonego neuronu, za pomocą której dokonywana jest klasyfikacja elementów.



Wykres 1. Położenie elementów klas w przestrzeni cech wraz z linią neuronu rozróżniającą klasy, z lewej dane uczące, z prawej dane testujące przy zestawie uczącym po 5 elementów

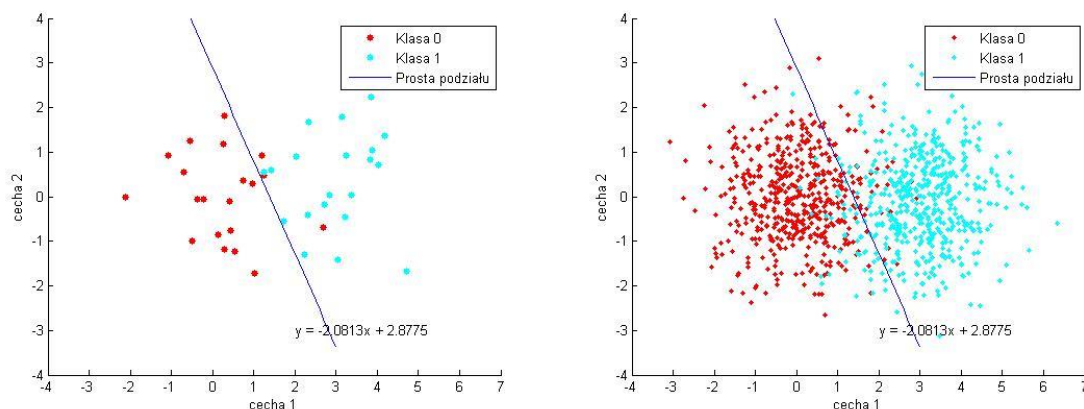
b) 20 elementów obu klas w zbiorze uczącym

Perceptron nauczony zestawem po 20 elementów z obu klas uzyskał następujące wyniki. Poniżej przedstawiono tabelę pomyłek klasyfikacji elementów zarówno dla zbioru uczącego jak i testującego. Neuron pomylił się w 7.5% dla danych uczących oraz w 9.1% dla danych testujących.

Tabela 2. Macierz pomyłek klasyfikacji perceptronu przy zestawie uczącym po 20 elementów

	Dane uczące zakwalifikowane do klasy 0	Dane uczące zakwalifikowane do klasy 1	Dane testowe zakwalifikowane do klasy 0	Dane testowe zakwalifikowane do klasy 1
Klasa 0	17	3	454	46
Klasa 1	0	20	45	455

Poniższy wykres przedstawia położenie elementów klas w przestrzeni cech wraz z linią wyuczonego neuronu, za pomocą której dokonywana jest klasyfikacja elementów.



Wykres 2. Położenie elementów klas w przestrzeni cech wraz z linią neuronu rozróżniającą klasy, z lewej dane uczące, z prawej dane testujące, przy zestawie uczącym po 20 elementów

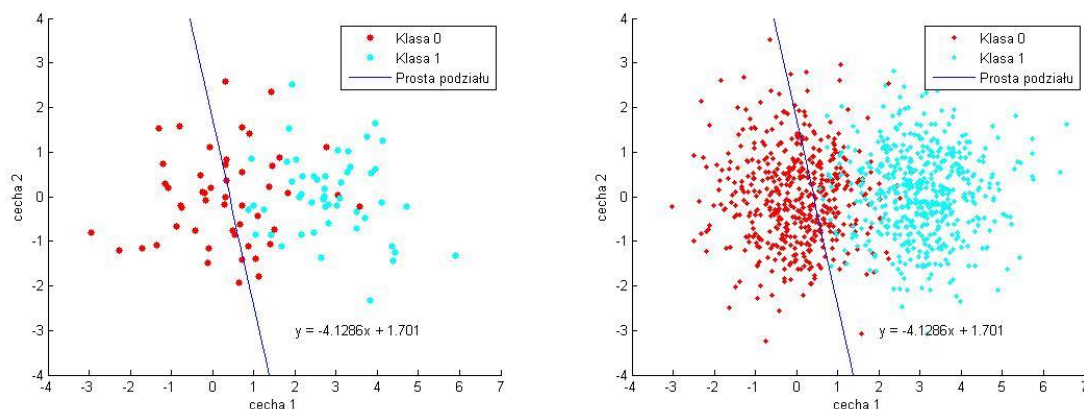
c) 50 elementów obu klas w zbiorze uczącym

Perceptron nauczony zestawem po 50 elementów z obu klas uzyskał następujące wyniki. Poniżej przedstawiono tabelę pomyłek klasyfikacji elementów zarówno dla zbioru uczącego jak i testującego. Neuron pomylił się w 23% dla danych uczących oraz w 15.8% dla danych testujących.

Tabela 3. Macierz pomyłek klasyfikacji perceptronu przy zestawie uczącym po 50 elementów

	Dane uczące zakwalifikowane do klasy 0	Dane uczące zakwalifikowane do klasy 1	Dane testowe zakwalifikowane do klasy 0	Dane testowe zakwalifikowane do klasy 1
Klasa 0	27	23	342	158
Klasa 1	0	50	0	500

Następnie sporządzono wykres przedstawiający położenie elementów klas w przestrzeni cech wraz z linią wyuczonego neuronu, za pomocą której dokonywana jest klasyfikacja elementów.



Wykres 3. Położenie elementów klas w przestrzeni cech wraz z linią neuronu rozróżniającą klasy, z lewej dane uczące, z prawej dane testujące, przy zestawie uczącym po 50 elementów

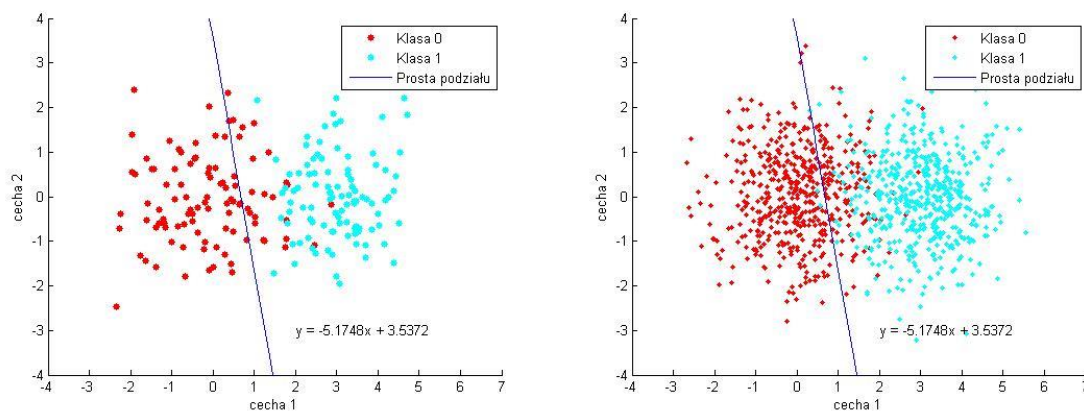
d) 100 elementów każdej klasy w zbiorze uczącym

Perceptron nauczony zestawem po 100 elementów z obu klas uzyskał następujące wyniki. Poniżej przedstawiono tabelę pomyłek klasyfikacji elementów zarówno dla zbioru uczącego jak i testującego. Neuron pomylił się w 12% dla danych uczących oraz w 12.9% dla danych testujących.

Tabela 4. Macierz pomyłek klasyfikacji perceptronu przy zestawie uczącym po 100 elementów

	Dane uczące zakwalifikowane do klasy 0	Dane uczące zakwalifikowane do klasy 1	Dane testowe zakwalifikowane do klasy 0	Dane testowe zakwalifikowane do klasy 1
Klasa 0	76	24	373	127
Klasa 1	0	100	2	498

Poniższy wykres przedstawia położenie elementów klas w przestrzeni cech wraz z linią wyuczonego neuronu, za pomocą której dokonywana jest klasyfikacja elementów.



Wykres 4. Położenie elementów klas w przestrzeni cech wraz z linią neuronu rozróżniającą klasy, z lewej dane uczące, z prawej dane testujące, przy zestawie uczącym po 100 elementów

Perceptron uzyskał najlepszą sprawność klasyfikacji zbioru testowego dla zbioru uczącego o wielkości 20 elementów obu klas i wynosi ona 90.9%. Uzyskane dane pokazują, że zwiększanie wielkości zbioru uczącego nie wpływa na poprawę sprawności perceptronu w klasyfikacji elementów. Wynika to z faktu, że zadany problem jest dla niego trudny. Elementy obu klas w pewnym stopniu przenikają się i w tym obszarze nie dają się odróżnić. Co więcej elementy obu klas rozróżnia tylko cecha pierwsza, ponieważ obie klasy mają taki sam rozkład cechy drugiej.

3. Realizacja ćwiczenia – dane rzeczywiste

Kolejnym ćwiczeniem było zastosowanie perceptronu do klasyfikacji próbek benzyny do jednej z dwóch klas czystości oznaczonych jako 0 oraz 1. Klasyfikacja została dokonana na podstawie trzech cech opisujących fizykochemiczne właściwości próbek. Etap nauczania został powtórzony 5-krotnie. W każdym procesie nauczania ilość epok – iteracji uaktualniania wartości wag sieci była taka sama i wynosiła 1000. Uzyskane wyniki po każdym uczeniu różnią się od siebie i zostały zestawione w poniższej tabeli.

Tabela 5. Zestawienie macierzy pomyłek dla pięciu prób nauczania

	Uczenie nr. 1		Uczenie nr 2.		Uczenie nr. 3		Uczenie nr. 4		Uczenie nr. 5	
\Zaklasyfikowana jako	Kl. 0	Kl. 1	Kl. 0	Kl. 1	Kl. 0	Kl. 1	Kl. 0	Kl. 1	Kl. 0	Kl. 1
Klasa 0	15	0	14	1	15	0	14	1	13	2
Klasa 1	1	14	1	14	4	11	5	10	1	14

Powyższa tabela pokazuje, że sprawność klasyfikacji perceptronu dla każdej próby uczenia jest różna pomimo, że zbiór uczący był identyczny. Dzieje się tak ponieważ wagi połączeń inicjalizowane są wartościami losowymi, następnie proces nauczania dopasowuje je szukając wartości optymalnych. Dlatego też w każdej z prób proces nauczania kończy się z innymi wartościami wag, które determinują działanie perceptronu. Neuron uzyskał najlepszą sprawność klasyfikacji dla pierwszego procesu nauczania równą 96.7%, natomiast najgorsza sprawność została uzyskana dla czwartej próby nauczania i wyniosła 80%.

4. Podsumowanie

W ramach ćwiczenia został zanalizowany pojedynczy neuron typu perceptron stanowiący podstawę sieci neuronowej. Neuron został wykorzystany w celu klasyfikacji danych wygenerowanych losowo a także danych rzeczywistych. Dla danych losowych przetestowano różne wielkości zbioru uczącego i nie zauważono poprawy sprawności klasyfikacji przy rosnącym zbiorze uczącym. Wynika to z faktu, że problem postawiony przed neuronem jest trudny, elementy obu klas przykrywają się i nie dają się jednoznacznie oddzielić funkcją liniową. Co więcej rozkład jednej z cech dla obu wygenerowanych klas jest identyczny. Dla danych rzeczywistych proces nauczania został powtórzony pięciokrotnie i zaobserwowano, że w każdej próbie nauczania perceptron daje inne wyniki, co jest związane z losową inicjalizacją wag neuronu. Badany neuron daje się wykorzystać do problemu klasyfikacji pomiędzy dwoma klasami.