

SISE - notatki pod egzamin

Dawid Gradowski
@puckmoment na dc

15.06.2025

Contents

1	Co to jest sztuczna inteligencja?	4
2	Co to jest zmienna lingwistyczna?	4
3	Czym jest funkcja aktywacji, jaka jest najczęściej stosowana?	5
4	Co to są sieci neuronowe, do czego służą?	5
5	Dlaczego neurony nazywamy neuronami?	5
6	Co to dane i informacje, różnice między nimi?	5
7	Wzór na zbiór rozmyty?	6
8	Gdzie był wykorzystany perceptron, co to perceptron?	6
9	Perceptron wielowarstwowy	6
10	Czym jest wyostrzanie?	7
11	Wykresy wyostrzenia?	8
12	Czym jest Adaline?	9
13	Uczenie Adaline	10
14	Czym jest Madaline?	10
15	Jaki wzór sugerował zamiast sigmoidalnej? – tangens hiperboliczny	11
16	Sieć Kohonena	11
16.1	Kohonen – winner takes all (wzór)	11
16.2	Modyfikacja na winner takes most (dopisanie * odległość od sąsiadów)	12
17	Przeuczenie w wielowarstwowym perceptronie?	12
18	Algorytm genetyczny	12
19	Operacje genetyczne	12

20 Czym jest ruletka?	12
20.1 Wzór na wycinek?	12
21 Gwiazda wejść/wyjść?	12
22 Uczenie nadzorowane?	12
23 Metoda gradientowa?	12
24 Notacja wektorowa(Diraca)?	12
25 Funkcja celu	12
26 Momentum i alpha	12
27 Gaz neuronowy	12

Wstęp

Notatki na podstawie pytań, które wpadły mi w ręce. Nie jestem autorem wszystkich tekstów jakie zostaną tu umieszczone.

1 Co to jest sztuczna inteligencja?

Jest to imitacja inteligencji człowieka realizowana przez maszyny, pozwalająca im na wykorzystanie wrodzonej i zdobytej wiedzy do rozwiązywania nowych problemów

2 Co to jest zmienna lingwistyczna?

Zmienna lingwistyczna to zmienna, której wartościami są słowa lub wyrażenia języka naturalnego, a nie konkretne liczby.

Weźmy na przykład zmienną lingwistyczną „temperatura”. Zamiast przyjmować wartości liczbowe (np. 20°C, 30°C), może przyjmować wartości takie jak:

- zimno
- chłodno
- ciepło
- gorąco

Każda z tych wartości lingwistycznych może być dalej modelowana za pomocą zbiorów rozmytych – np. „ciepło” może obejmować temperatury od 20°C do 30°C, ale z różnym stopniem przynależności (np. 22°C to „trochę ciepło”, 28°C to „bardzo ciepło”).

3 Czym jest funkcja aktywacji, jaka jest najczęściej stosowana?

Funkcja aktywacji to pojęcie używane w sztucznej inteligencji do określenia funkcji, według której obliczana jest wartość wyjścia neuronów sieci neuronowej.

Najczęściej używana jest sigmoida:

$$y = \frac{1}{1 + \exp(-\beta s)}$$

w której

- β - wybiera użytkownik, wpływa na kształt wykresy, wzrost $\beta \rightarrow$ większy kąt
- s - sygnał

4 Co to są sieci neuronowe, do czego służą?

Sieciami neuronowymi nazywamy system którego budowa i zasada działania w pewnym stopniu wzorowana jest na funkcjonowaniu fragmentów rzeczywistego systemu nerwowego. Na przesłankach biologicznych oparte są schematy sztucznych neuronów wchodzących w skład sieci oraz jej struktura.

Służą do uczenia maszynowego, by maszyna mogła rozwiązywać nowe problemy, których nigdy nie widziała na podstawie wiedzy zdobytej podczas uczenia.

5 Dlaczego neurony nazywamy neuronami?

Sztuczne neurony które są w sieciach neuronowych są zbudowane na podobieństwo neuronów jako komórek nerwowych. Mają wejścia, coś co przetwarza i wyjścia.

6 Co to dane i informacje, różnice między nimi?

Dane zawsze reprezentują fakty, bez odpowiedniej interpretacji nic nie znaczą. Dopiero informacja co oznaczają te dane pozwala nam zrozumieć ich sens.

7 Wzór na zbiór rozmyty?

Zbiór par składający się z poszczególnych elementów przestrzeni rozważań i ich stopni przynależności do tego zbioru rozmytego.

Zbiory rozmyte (fuzzy) zawierają obiekty spełniające nieprecyzyjne własności przynależności; innymi słowy przynależność może być w pewien sposób przybliżona.

Zbiór rozmyty A w pewnej przestrzeni rozważań $X = x$ określa się jako zbiór par

$$A = (x, \mu_A(x))$$

gdzie

$\mu_A : X \rightarrow [0, 1]$ jest funkcją przynależności zbioru rozmytego A , a wartość $\mu_A \in [0, 1]$ jest stopniem przynależności elementu $x \in X$ do zbioru rozmytego A

Znaczy to mniej więcej tyle, że zamiast przynależności pełnej "tak"(1) i "nie"(0) jak w zbiorach klasycznych, mogą być to wartości stopnia przynależności od 0 do 1.

8 Gdzie był wykorzystany perceptron, co to perceptron?

Jest to najprostsza sieć neuronowa, składająca się z jednego bądź wielu niezależnych neuronów. Przeznaczeniem było rozpoznawanie znaków alfanumerycznych. Innowacją było tu zastosowanie procesu uczenia jako metody programowania systemu.

9 Perceptron wielowarstwowy

Może mieć 3 warstwy:

- Wejściową (linową),
- ukrytą (nieliniową),
- wyjściową (liniową).

Może mieć więcej warstw ale nie ma to sensu.

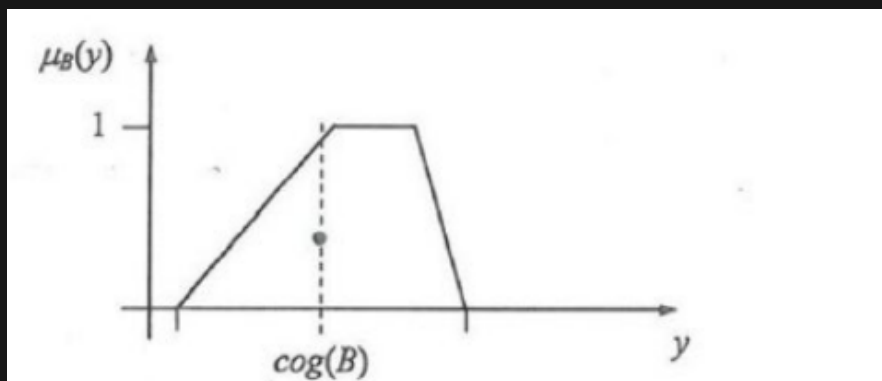
10 Czym jest wyostrzanie?

wyostrzanie (defuzzification) jest przekształceniem, dzięki któremu zbiór rozmyty B zostaje zastąpiony przez pojedynczą liczbę $\bar{y} \in Y$ będącą jego reprezentacją. Intuicja podpowiada, że wyostrzanie powinno być rodzajem uśredniania. Oto kilka popularnych metod.

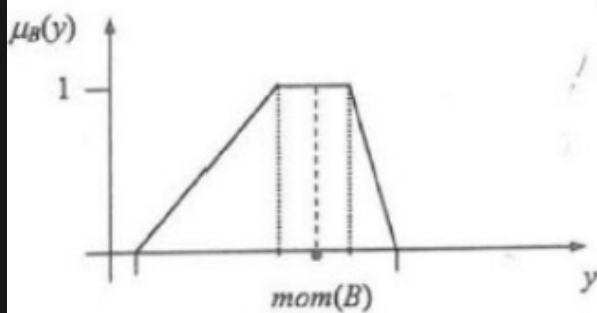
Metoda środka ciężkości (cog - center of gravity) jest znanym z fizyki rachunkiem, w którym masa (lub ciężar) jest zastępowana przez wartości funkcji przynależności. Dla ciągłych funkcji przynależności mamy:

$$\bar{y} = cog(B) = \frac{\int_A \mu_B(y) y \cdot dy}{\int_A \mu_B(y) \cdot dy}$$

11 Wykresy wyostrzenia?



Rys.3. Ilustracja wyznaczenia cog w przypadku ciągłym



Rys.4. Ilustracja wyznaczenia mom

12 Czym jest Adaline?

Adaline jest to adaptacyjny liniowy neuron (Adaptive Linear Neuron). Jego wagi mogą ulegać zmianie. Ma generować odpowiednie wyjście dla określonych wejść.

$$o = \sum_{j=1}^n w_j x_j = w^T x$$

Nauka odbywa się w procesie nadzorowanym z tzw. nauczycielem. Od wymaganej odpowiedzi r (właściwe rozwiązanie) odejmujemy o (wyjście) i otrzymujemy d . Otrzymana wartość potrzebna jest w regule Delta - czyli zmiany wagi w oparciu o d .

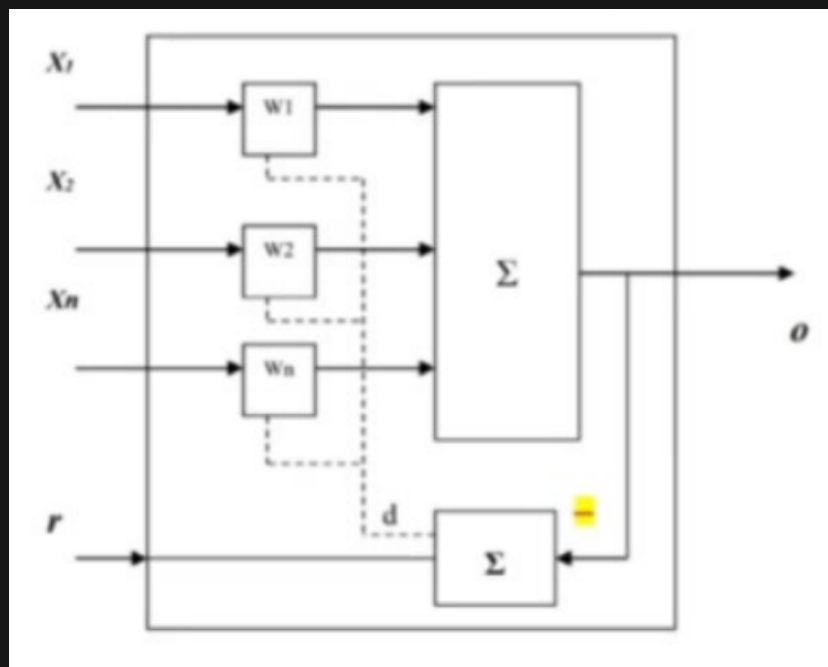
Reguła Delta (Δ):

$$w^{k+1} = w^k + \eta dx$$

$$d = r - o$$

gdzie

- r - wartość oczekiwana
- o - aktualna wartość na wyjściu



13 Uczenie Adaline

Uczenie Adaline jest oparte o Regułę Delta (Δ):

$$w^{k+1} = w^k + \eta dx$$

$$d = r - o$$

gdzie

- r - wartość oczekiwana (cel)
- o - aktualna wartość na wyjściu

Interpretacja tej reguły:

Niech $d > 0$, tzn. $r > o$. Oznacza to, że sygnał wyjściowy z neuronu jest za mały.

sygnał wyjściowy zależy od kąta pomiędzy wektorami x i w ; kąt jest za duży.

$$o = \sum w_i x_i = w^T x \Leftarrow \text{iloczyn skalarny}$$

Jeśli x i w są znormalizowane, tzn. $\|x\| = x^T x = \|w\| = w^T w = 1$, to $o = \cos \alpha$, gdzie kąt α jest kątem między wektorami x i w .

Należy uzgodnić kierunki x i w . Dodając wektorowo x do wektora w , otrzymuje się nowy wektor w^{k+1} bardziej zgodny z r niż poprzedni.

Ponieważ zwykle $\eta d < 1$, to wg wzoru (Δ) dodaje się tylko fragment wektora x , co zapobiega zbyt gwałtownym obrotom wektora w .

Gdy $d < 0$, tzn. $r < o$, następuje oddalanie wektorów x i w ; odpowiedź była zbyt silna.

14 Czym jest Madaline?

Madaline (Multiple ADaptive LINear Elements) to sieć neuronowa zbudowana z wielu adaptacyjnych neuronów liniowych (ADALINE), które tworzą jedną lub więcej warstw. Każdy neuron otrzymuje wszystkie dane wejściowe i dokonuje liniowej kombinacji wag oraz sygnałów wejściowych, a następnie stosuje funkcję progową. Dzięki odpowiedniej organizacji neuronów i zasadzie głosowania, sieć może rozpoznawać także funkcje nieliniowe. W niektórych przypadkach struktura może zawierać więcej neuronów niż jest to konieczne do realizacji danego zadania, co czyni ją potencjalnie nadmiarową.

15 Jaki wzór sugerował zamiast sigmoidalnej? – tangens hiperboliczny

Tangens hiperboliczny:

$$\operatorname{tgh} x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

może być również oznaczone jako $\tanh x$ i $\operatorname{th} x$.

16 Sieć Kohonena

Jest to najbardziej znana i najczęściej stosowana sieć samoucząca się, realizuje zasadę samoorganizacji (SOM). Jest to także najbardziej znany przykład sieci konkurencyjnej wykorzystującej koncepcję sąsiedztwa. W wyniku uczenia tej sieci powstaje mapa topologiczna, której aprioryczna interpretacja jest niemożliwa (bo sieć uczy się bez nauczyciela i użytkownik nie ma kontroli nad tym co sieć robi). Jednak po nauczaniu można zwykle ustalić, jakie znaczenie mają poszczególne rejony tej mapy (tworzonej przez sygnały wyjściowe pochodzące z warstwy topologicznej) na podstawie analizy konkretnych danych wejściowych.

Klasycznie: Sieć wybiera neuron o najwyższym poziomie aktywacji i w ten sposób oeklaruje go jako zwycięzce. Zwycięzka bierze wszystko (w literaturze ang.: winner takes all).

Tym samym sieć jest **klasyfikatorem**.

Liczba neuronów liniowych wyznacza liczbę klas, które potencjalnie sieć może rozróżniać. Elementami wektora wejść x są cechy na podstawie których rozróżnia się klasy.

CEL NAUKI: $w_i^{p*(k+1)} \sim x_i^{(k)}$

Dodatkowe pytania:

- Ile ma warstw kohonen? (1)
- Może mieć więcej? (nie)
- Dlaczego? (bo założenie funkcji liniowych nadal jest liniowe wiec nie ma sensu)

16.1 Kohonen – winner takes all (wzór)

$$w_i^{p(k+1)} = w_i^{p(k)} + \eta^{(k)}(x_i^{(k)} - w_i^{p(k)})$$

gdzie:

- w - wektor wag
- x - wejście
- η - współczynnik kroku
- p - pokazuje, że to zwycięzca

16.2 Modyfikacja na winner takes most (dopisanie * odległość od sąsiadów)

Chuj wie :)

17 Przeuczenie w wielowarstwowym perceptronie?

18 Algorytm genetyczny

19 Operacje genetyczne

20 Czym jest ruletka?

20.1 Wzór na wycinek?

21 Gwiazda wejść/wyjść?

22 Uczenie nadzorowane?

23 Metoda gradientowa?

24 Notacja wektorowa(Diraca)?

25 Funkcja celu

26 Momentum i alpha

27 Gaz neuronowy