

Sieci neuronowe, wykład 8

Joanna Jędrzejowicz

Instytut Informatyki

Dlaczego sieci neuronowe

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

- inspiracja naturą - skomplikowane systemy uczące się składają się z połączonych między sobą neuronów,
- rzeczywisty neuron ma dendryty zbierające sygnały pochodzące od innych neuronów; neuron przetwarza sygnały wejściowe tworząc wyjście kierowane do następnych neuronów poprzez akson i synapsy pełniące funkcję przekaźników do innych neuronów; sygnały mogą zostać wzmocnione lub osłabione
- neuron sumuje impulsy pobudzające i hamujące i jeśli suma przekracza pewną wartość progową, to sygnał wyjściowy jest przesyłany do innych neuronów,

Sformalizowany model neuronu

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Teoretyczny model neuronu:

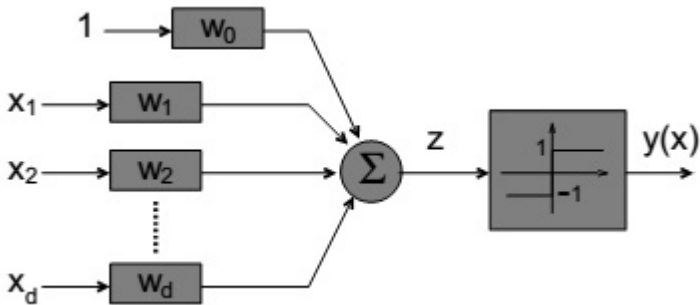
- d - liczba wejść w neuronie,
- x_1, \dots, x_d - sygnały wejściowe,
- w_1, \dots, w_d - wagi synaptyczne,
- w_0 wartość progowa (po angielsku bias), f funkcja aktywacji
- $y = f(s)$, gdzie $s = \sum_{i=1}^d w_i * x_i + w_0$

Sieć trzeba nauczyć! Uczenie polega na modyfikacji wag przy wykorzystaniu danych treningowych.

Perceptron

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz



funkcja aktywacji jest tutaj bipolarna:

$$f(x) = \begin{cases} 1 & \text{gdy } x > 0 \\ -1 & \text{inaczej} \end{cases}$$

Inne funkcje aktywacji

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Funkcja sigmoidalna

$$f(x) = \frac{1}{1 + \exp(-x)}$$

Funkcja skokowa

$$f_{\theta}(x) = \begin{cases} 1 & \text{gdy } x \geq \theta \\ 0 & \text{inaczej} \end{cases}$$

Perceptron

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Perceptron może przyjmować dwie wartości, czyli klasyfikować sygnały wejściowe x_1, \dots, x_d do dwóch klas

Dla $d = 2$ perceptron dzieli płaszczyznę na dwie części, podział wyznacza prosta o równaniu

$$w_1 * x_1 + w_2 * x_2 + w_0 = 0$$

Prosta wyznaczająca podział przestrzeni jest prostopadła do wektora $[w_1, w_2]$.

Uczenie perceptrona - uczenie nadzorowane

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Uczenie ma sens dla zbiorów danych liniowo separowalnych

- wybieramy losowo wagi początkowe oraz próg
- iterujemy następujące postępowanie
 - dla kolejnego wektora uczącego $x = [x_1, \dots, x_n]$ z wartością wyjściową y_d oblicz wartość wyjścia sieci y
 - jeżeli $y = y_d$, to wagi pozostają bez zmian,
 - jeżeli $y \neq y_d$, to wagi modyfikuje się: $e = y_d - y$,
 $w_i \leftarrow w_i + \alpha * e * x_i$ (α jest zadany współczynnikiem wzrostu z przedziału $(0, 1)$); jeżeli błąd e jest dodatni, to trzeba zwiększyć wartość y , każde wejście dostarcza $x_i * w_i$, czyli jeżeli x_i dodatnie, to wzrost w_i zwiększa y , jeśli ujemne to zmniejsza y ; analogicznie jeśli błąd ujemny
- algorytm powtarza się aż dla wszystkich wektorów wejściowych błąd będzie mniejszy od założonej tolerancji, bądź wartości wag ustabilizują się

przykładSieci.xls

Twierdzenie o zbieżności

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Jeżeli zbiór danych treningowych jest **liniowo separowalny** oraz istnieje zestaw wag klasyfikujący dane treningowe poprawnie tzn. wyznaczający odwzorowanie $y = f(x)$, to algorytm uczący znajdzie rozwiązanie w skończonej liczbie iteracji dla dowolnych wartości początkowych wektora wag.

Przykład

Sieci
neuronowe,
wykład 8

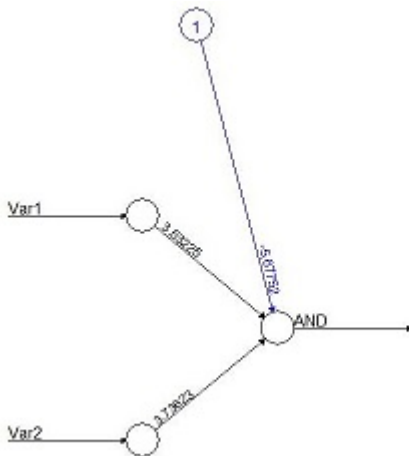
Joanna
Jędrzejowicz

```
# siec wyliczajaca AND
library("neuralnet")
AND <- c(rep(0,3),1)
bin.data <- data.frame(expand.grid(c(0,1), c(0,1)), AND)
print(net.and <- neuralnet(AND~Var1+Var2,
bin.data, hidden=0, rep=10, linear.output=FALSE))
print(net.and$weights)
plot(net.and, rep="best")
```

Przykład, perceptron

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz



wagi:

$w_0 = -5.561084069$, $w_1 = 3.620697009$, $w_2 = 3.611874687$ z
zastosowaniem fkcji sigmoidalnej

Przykład - zbiór danych nie jest liniowo separowalny

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Funkcja XOR:

| x | y | XOR |
|---|---|-----|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |

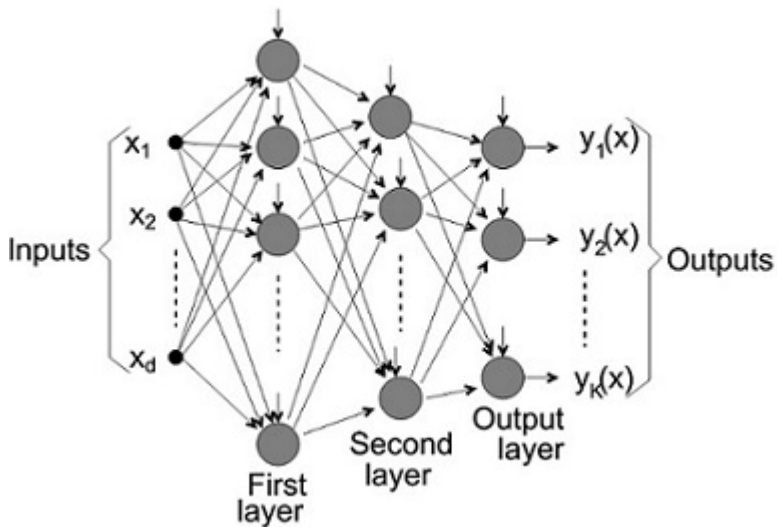
Jak wygląda wykres?

Trzeba zastosować **sieć** neuronową wielowarstwową

Sieć

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz



Sieć neuronowa

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Sieć jest charakteryzowana przez:

- architekturę sieci - położenie poszczególnych neuronów i powiązania między nimi,
- funkcję aktywacji,
- metodę trenowania (uczenia) sieci

Sieć neuronowa

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

- w sieci neuronowej jednokierunkowej przepływ jest dozwolony tylko w jednym kierunku, nie ma pętli i cykliów,
- sieć może się składać z wielu warstw, ale najczęściej stosuje się warstwę wejściową, ukrytą i wyjściową,
- liczba neuronów w warstwie wejściowej zależy od liczby atrybutów w zbiorze danych,
- duża liczba neuronów w warstwie ukrytej powoduje zwiększenie mocy obliczeniowej i ułatwia rozpoznawanie skomplikowanych wzorców, ale może prowadzić do przeuczenia sieci

przykład sieć wyliczająca pierwiastek sqrt.jpeg

Sieć wielowarstwowa

Sieci
neuronowe,
wykład 8

Joanna
Jędrzejowicz

Każda warstwa ma swoje zadania:

- warstwa wejściowa akceptuje sygnały wejściowe (ze świata zewnętrznego) i dokonuje redystrybucji do neuronów w warstwie ukrytej,
- warstwa wyjściowa ustala sygnały wyjściowe pobierając wyjście z warstwy ukrytej,
- wyjścia z warstwy ukrytej nie mogą być zaobserwowane, stąd nazwa.