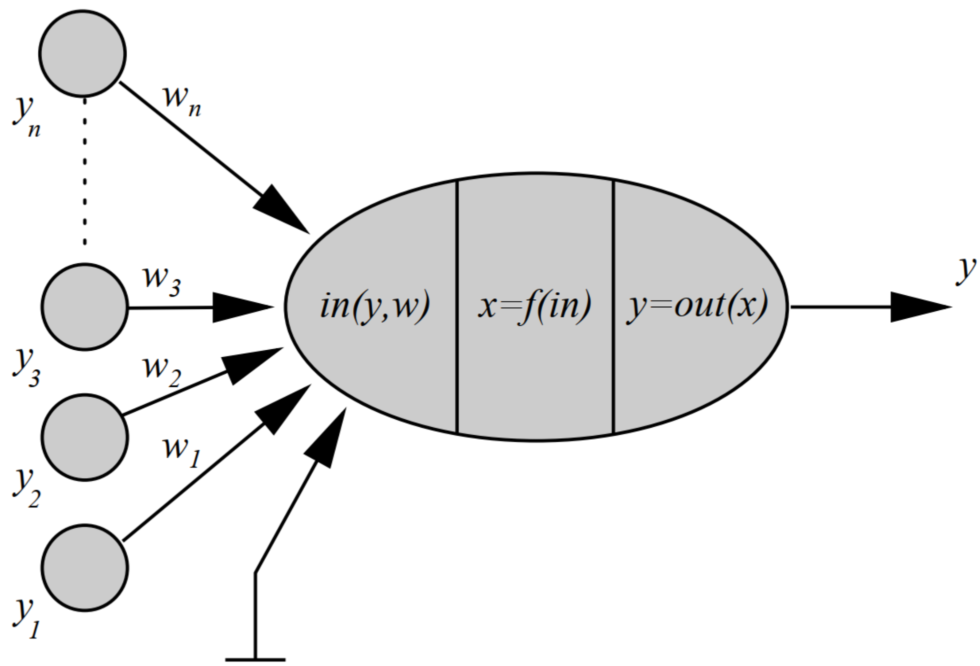
Cvičenie 1 – Matematické základy neurónových sietí

Pred tým než sa spustíme do štúdia neurónových sietí, na dnešnom cvičení si zopakujeme niektoré matematické operácie, ktoré sa nám zídu pri pochopení rôznych konceptov v neurónových sieťach.

Ako už poznáte z predmetov Základy inteligencie systémov a Umelá inteligencia, základným stavebným blokom neurónových sietí je neurón, ktorý je zároveň procesnou jednotkou. Jej štruktúru môžeme znázorniť nasledovne:



Skalárny súčin

Jedna zo základných operácií, ktoré sa odohrávajú v neuróne je výpočet váženého súčtu vstupov. V skutočnosti sa jedná o prenásobenie každej vstupnej hodnoty s príslušnou váhou a o následný výpočet súčtu týchto násobkov (s pripočítaním prahu).

Ak teda vstupy a váhy reprezentujeme ako jednodimenzionálne vektory, naším cieľom je vypočítať skalárny súčin dvoch vektorov:

$$X = (x_1, x_2, \dots, x_n)$$
$$W = (w_1, w_2, \dots, w_n)$$
$$X \cdot W = x_1 w_1 + x_2 w_2 + \dots + x_n w_n$$

V niektorých učebniciach prah neurónu je reprezentovaný ako nultá váha so vstupom 1.

Precvičte si výpočet skalárneho súčtu na nasledujúcich príkladoch:

$$A = (4, -2, 0, 3, 1)$$
$$B = (-1, 1, 2, 5, 6)$$
$$C = (-3, -1, 0, 1, 1)$$
$$D = (1, 7, -2, 1, 2)$$
$$E = (0, 0, 2, 3, 1)$$

Vypočítajte skalárne súčiny pre:

$$A \cdot C$$
$$D \cdot E$$
$$A \cdot D$$
$$B \cdot C$$
$$E \cdot B$$

Násobenie matíc

V posledných rokoch grafické karty sú čoraz zaužívanejšie pre trénovanie hlbokých neurónových sietí pretože sú veľmi vhodné pre maticové operácie. Pre trénovanie neurónových sietí matice sú používane pri výpočte zmien váh a pri samotnej aktualizácii hodnôt váh.

Matice dokážeme násobiť jedine v prípade ak počet stĺpcov prvej matice je rovnaký ako počet riadkov druhej matice. Maticový produkt má potom rozmery počet riadkov ako v prvej matici krát počet stĺpcov ako v druhej matici. Výsledná hodnota na každej pozícii produktu sa vypočíta nasledovne:

$$(AB)[i, j] = A[i, 1]B[1, j] + A[i, 2]B[2, j] + \cdots + A[i, n]B[m, j]$$

Majme matice:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 4 & 0 \end{pmatrix}$$

$$B = \begin{pmatrix} 3 & 1 \\ 2 & 2 \\ 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} 5 & 7 \\ 1 & 3 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 5 & 0 \\ -1 & 0 & -1 \end{pmatrix}$$

Vypočítajte všetky možné maticové produkty!

Derivácia

Derivácia funkcie predstavuje zmenu tejto funkcie pri malej zmene vstupných premenných. Pri riešení neurónových sietí sa často stretávame s parciálnou deriváciou, čo je derivácia funkcie s viacerými premennými podľa jednej premennej (ostatné premenné sa považujú za konštanty).

Pravidlá pre deriváciu najčastejších funkcií sú nasledovné:

$$f(x) = c \to f'(x) = 0$$
$$f(x) = x^n \to f'(x) = nx^{n-1}$$
$$f(x) = n^x \to f'(x) = n^x \ln n$$
$$f(x) = e^x \to f'(x) = e^x$$
$$f(x) = log_a x \to f'(x) = \frac{1}{x \cdot \ln a}$$
$$f(x) = \sin x \to f'(x) = \cos x$$
$$f(x) = \cos x \to f'(x) = -\sin x$$

Ak chcete naraz derivovať niekoľko funkcií, platia nasledujúce zásady:

$$(f+g)' = f' + g'$$
$$(fg)' = f'g + fg'$$
$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$
$$f\big(g(x)\big)' = f'\big(g(x)\big) \cdot g'(x)$$

Určte deriváciu nasledujúcich funkcií:

$$x^{10} + 3x^5 + 2x^3 + 4$$
$$e^2 + 1$$
$$\frac{x}{2} + \pi^3$$
$$(x^3 + x)(x^5 + x^2)$$
$$(x^2 + 2)^2$$
$$(x^{10}(x^2 + 1)^{10})$$
$$\frac{5}{(x+5)^2}$$
$$\frac{x^2 + 5x}{x + 5}$$

Na záver zderivujte sigmoidálnu funkciu, ktorá sa často používa ako aktivačná funkcia v neurónoch:

$$\frac{1}{1 + e^{-x}}$$