

## Temat: Rola funkcji aktywacji w modelu neuronu oraz typy najczęściej stosowanych funkcji aktywacji

Sieci neuronowe, będące fundamentalnym elementem współczesnej sztucznej inteligencji i uczenia maszynowego, zawdzięczają swoją potęgę zdolności modelowania nawet skomplikowanych, nieliniowych zależności. Kluczową rolę w tym procesie odgrywają funkcje aktywacji.

W kontekście modelu neuronowego, funkcje aktywacji są integralną częścią architektury sieci. Każdy neuron przyjmuje wiele sygnałów wejściowych, które są ważone przez odpowiadające im wagi. Suma iloczynów tych sygnałów i wag przepływa przez funkcję aktywacji, której zadaniem jest decyzja, czy i w jakim stopniu neuron powinien aktywować się i przekazywać informacje dalej do kolejnej warstwy sieci.

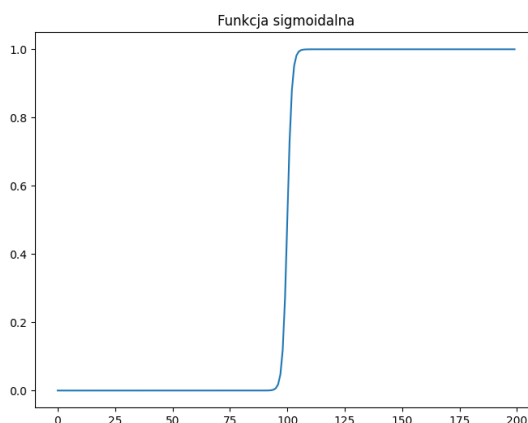
Funkcje aktywacji wprowadzają do modelu niezbędne nieliniowości. Bez nich, mimo wielu warstw ukrytych, cały model byłby wciąż jedynie liniowym transformatorów wejść, co ograniczałoby zdolność sieci do efektywnego uczenia się złożonych wzorców danych. Dzięki zastosowaniu funkcji aktywacji, sieć neuronowa zyskuje zdolność do modelowania skomplikowanych, nieliniowych zależności między danymi.

Możemy wyróżnić wiele rodzajów funkcji aktywacji, z których każda ma swoje unikalne właściwości i jest najbardziej odpowiednia dla określonych zastosowań. Popularne funkcje aktywacji przedstawię krótko poniżej.

Funkcja sigmoidalna, będąca jedną z najstarszych funkcji aktywacji, mapuje wejście do zakresu między 0 a 1. Choć jest gładka i różniczkowalna, co jest istotne dla procesu uczenia sieci, ma wady takie jak zanikający gradient, który może spowolnić proces uczenia.

$$\sigma(x) = \frac{1}{1+e^{-x}}, \text{ gdzie}$$

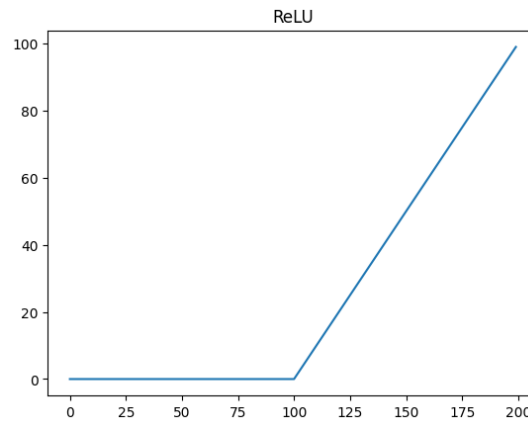
$\sigma(x)$  – wartość funkcji sigmoidalnej dla wejścia  $x$



Funkcja tangensu hiperbolicznego (tanh) mapuje wejście do zakresu między -1 a 1. Jest to często lepsze niż zakres od 0 do 1, ponieważ średnia wartość aktywacji w sieci jest bliższa zero.

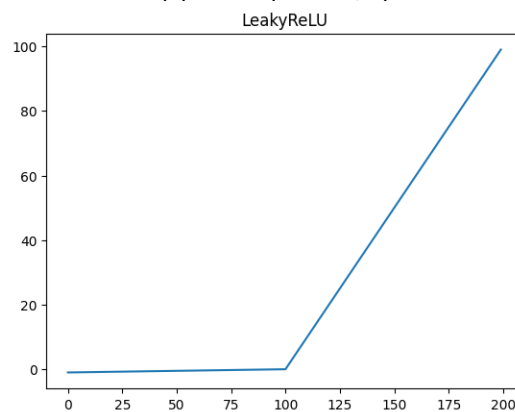
Funkcja Rectified Linear Unit (ReLU) stała się bardzo popularna z powodu swojej prostoty i efektywności. ReLU daje wynik równy zero dla wartości wejściowych mniejszych od zera, a dla wartości dodatnich jest tożsama z wejściem. ReLU radzi sobie lepiej z problemem zanikającego gradientu, ale wprowadza problem "martwych neuronów", które nigdy nie są aktywowane.

$$\sigma(x) = \max(0, x)$$



Leaky ReLU i Parametric ReLU to modyfikacje funkcji ReLU, które próbują zaradzić problemowi martwych neuronów, wprowadzając mały gradient dla wartości wejściowych mniejszych od zera. Przykład dla Leaky ReLU:

$$\sigma(x) = \max(0.01 * x, x)$$

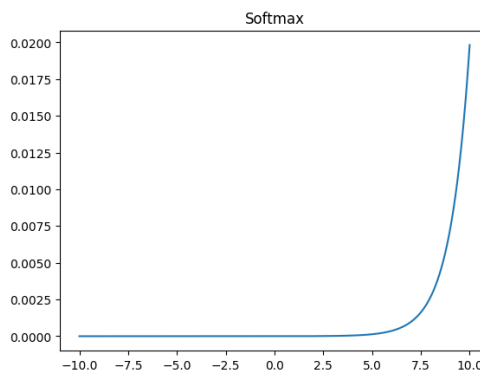


Funkcja Softmax jest często używana w ostatniej warstwie sieci neuronowych dla zadań klasyfikacji wieloklasowej. Softmax przekształca wejścia w postaci logitów na prawdopodobieństwa sumujące się do jedności, co umożliwia interpretację wyników jako prawdopodobieństwa przynależności do poszczególnych klas.

$$\sigma(e_i) = \frac{e^{x_j}}{E(e^{x_j})}, \text{ gdzie:}$$

$x_i$  – i-ty element wejściowego wektora  $x$ ,

$E(e^{x_j})$  – suma wszystkich  $e^{x_j}$ , gdzie  $x_j$  oznacza każdy element wektora wejściowego  $x$ .



Podsumowując, funkcje aktywacji pełnią kluczową rolę w sieciach neuronowych, wprowadzając niezbędne nieliniowości do modelu, które pozwalają na efektywne modelowanie złożonych zależności. Bez funkcji aktywacji, sieci neuronowe byłyby jedynie liniowymi transformatorem wejść, co znacznie ograniczałoby ich zdolność do rozwiązywania rzeczywistych, często nieliniowych problemów. Wybór odpowiedniej funkcji aktywacji jest kluczowy dla projektowania efektywnych i wydajnych sieci neuronowych.

Opracował:

Dawid Walicki