Differentiëren

Vragen en antwoortden

Wanneer is een relatie een functie?
Als er 1 output is (Er kunnen meerdere inputs zijn)

Wat is differentiëren en waarom doe je het?
Differentiëren is het vinden van de veranderingssnelheid van een functie. In AI gebruiken we differentiatie om de afgeleide van een functie te berekenen. Dit is belangrijk voor machine learning en deep learning, waar we modellen willen optimaliseren. Het helpt algoritmen om automatisch de parameters aan te passen en de prestaties te verbeteren. Door differentiatie begrijpen we hoe veranderingen de modellen beïnvloeden en kunnen we optimalisatieproblemen oplossen.

Wat is een activatie functie?
Activatie Functies zijn wiskundige functies die op de output van neuronen worden toegepast om een activatiegetal te bepalen.

Waarom zijn afgeleide functies belangrijk bij het maken van een AI?
1. Optimalisatie: Ze sturen optimalisatie algoritmen, zoals gradient descent, om modellen te verbeteren door parameters aan te passen.

2. Backpropagation**:** Bij het trainen van neurale netwerken helpen afgeleide functies bij het berekenen van de aanpassingen van gewichten om fouten te minimaliseren.

3. Functie Optimalisatie: Ze helpen bij het vinden van optimale instellingen, zoals hyperparameters, voor AI-modellen.

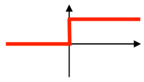
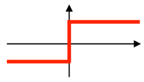
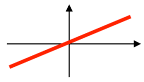
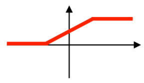
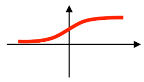
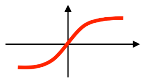
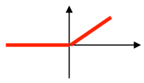
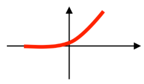
Kort gezegd: afgeleide functies sturen het proces van leren en verbeteren in AI.

Hoe bereken je de steilheid van een lijn?
Delta y / Delta x (Delta is het verschil tussen twee punten)

Regels bij het differentiëren

$f(x) = c$	$f'(x) = 0$
$f(x) = ax + b$	$f'(x) = a$
$f(x) = x^2$	$f'(x) = 2x$
$f(x) = x^n$	$f'(x) = nx^{n-1}$
$f(x) = \sqrt{x}$	$f'(x) = \frac{1}{2\sqrt{x}}$
$f(x) = \frac{1}{x^n}$	$f'(x) = -\frac{n}{x^{n+1}}$

Verschillende activatie functies

Activation function	Equation	Example	1D Graph
Unit step (Heaviside)	$\phi(z) = \begin{cases} 0, & z < 0, \\ 0.5, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Sign (Signum)	$\phi(z) = \begin{cases} -1, & z < 0, \\ 0, & z = 0, \\ 1, & z > 0, \end{cases}$	Perceptron variant	
Linear	$\phi(z) = z$	Adaline, linear regression	
Piece-wise linear	$\phi(z) = \begin{cases} 1, & z \geq \frac{1}{2}, \\ z + \frac{1}{2}, & -\frac{1}{2} < z < \frac{1}{2}, \\ 0, & z \leq -\frac{1}{2}, \end{cases}$	Support vector machine	
Logistic (sigmoid)	$\phi(z) = \frac{1}{1 + e^{-z}}$	Logistic regression, Multi-layer NN	
Hyperbolic tangent	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	Multi-layer Neural Networks	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = max(0, z)$	Multi-layer Neural Networks	
Rectifier, softplus 	$\phi(z) = \ln(1 + e^z)$	Multi-layer Neural Networks	

Verschillende rekenregels

Somregel: De afgeleide van de som van twee functies is gelijk aan de som van de afgeleiden van die functies.
(f + g)' = f' + g'

Productregel: De afgeleide van het product van twee functies is gelijk aan de afgeleide van de eerste functie keer de tweede functie plus de eerste functie keer de afgeleide van de tweede functie.
(f * g)' = f'g + fg'

Kettingregel: De afgeleide van een samengestelde functie is gelijk aan de afgeleide van de buitenste functie, geëvalueerd op de binnenste functie, vermenigvuldigd met de afgeleide van de binnenste functie.
(f(g(x))' = f'(g(x)) * g'(x)

Rekenvoorbeelden

Helling bepalen

1. **Kies een functie**:
Bijvoorbeeld, (f(x) = x^2).

2. **Bereken de afgeleide**:
De afgeleide van (f(x)) is (f'(x) = 2x).

3. **Kies een punt**:
Laten we zeggen dat we de helling van (f(x)) op (x = 3) willen vinden.

4. **Substitueer het punt in de afgeleide**:
Voor (x = 3), wordt de afgeleide (f'(3) = 2 keer 3 = 6)

5. **Interpreteer de helling**
Bij (x = 3), is de helling van de functie (f(x) = x^2) gelijk aan 6.

Activatie functies differentiëren

Sigmoid Functie (Logistische Functie):

De sigmoid functie wordt gegeven door:

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

De afgeleide van de sigmoid functie wordt als volgt berekend:

$$\frac{d\sigma(x)}{dx} = \frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right)$$

We gebruiken de quotiëntenregel om de afgeleide te berekenen:

$$= -\frac{1}{(1+e^{-x})^2} \cdot \frac{d}{dx}(1 + e^{-x})$$

$$= -\frac{1}{(1+e^{-x})^2} \cdot e^{-x} \cdot \frac{d}{dx}(-x)$$

$$= -\frac{e^{-x}}{(1+e^{-x})^2} \cdot (-1)$$

$$= \frac{e^{-x}}{(1+e^{-x})^2}$$

Rectified Linear Unit (ReLU) Functie:

De ReLU-functie is gedefinieerd als:

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

De afgeleide van de ReLU-functie is:

$$f'(x) = \begin{cases} 0 & \text{als } x < 0 \\ 1 & \text{als } x > 0 \end{cases}$$

Voor $x = 0$ is de afgeleide niet gedefinieerd.

Hyperbolische tangensfunctie Tanh

Om de afgeleide van de hyperbolische tangensfunctie (tanh) te vinden, gebruiken we de kettingregel. De hyperbolische tangensfunctie kan worden geschreven als:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Laten we deze functie differentiëren:

$$\frac{d}{dx}\tanh(x) = \frac{d}{dx}\left(\frac{e^x - e^{-x}}{e^x + e^{-x}}\right)$$

We kunnen de quotiëntregel toepassen:

$$= \frac{(e^x + e^{-x})(e^x + e^{-x}) - (e^x - e^{-x})(e^x - e^{-x})}{(e^x + e^{-x})^2}$$

$$= \frac{(e^{2x} + 2 + e^{-2x}) - (e^{2x} - 2 + e^{-2x})}{(e^x + e^{-x})^2}$$

$$= \frac{4}{(e^x + e^{-x})^2}$$

Dus, de afgeleide van de hyperbolische tangensfunctie is:

$$\frac{d}{dx}\tanh(x) = \frac{4}{(e^x + e^{-x})^2}$$

Partiële differentiatie

Laten we een voorbeeld bekijken waarbij we de partiële afgeleide berekenen van een functie $f(x, y)$ met betrekking tot x.

Stel dat we de functie hebben:

$$f(x, y) = x^2 + 3xy + y^2$$

Om de partiële afgeleide $\frac{\partial f}{\partial x}$ te vinden, differentiëren we de functie $f(x, y)$ naar x, waarbij y constant wordt gehouden.

Dus, we differentiëren $f(x, y)$ met betrekking tot x:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^2 + 3xy + y^2)$$

$$= \frac{\partial}{\partial x}(x^2) + \frac{\partial}{\partial x}(3xy) + \frac{\partial}{\partial x}(y^2)$$

$$= 2x + 3y$$

Dus, $\frac{\partial f}{\partial x} = 2x + 3y$.

Differentiëren van een matrice

Stel dat we een matrixfunctie hebben $F(X)$, waarbij X een matrix is en F een matrix naar matrix-functie. Bijvoorbeeld:

$$F(X) = X^2$$

Hier differentiëren we de matrix X elementgewijs (dat wil zeggen, elk element van de matrix afzonderlijk).

Om de afgeleide van $F(X)$ te vinden, differentiëren we elk element van de matrix X afzonderlijk. Dus, als $X = [x_{ij}]$, waarbij x_{ij} het element in de i-de rij en j-de kolom van X is, dan is de afgeleide van $F(X)$ (genoteerd als $\frac{dF(X)}{dX}$) gelijk aan:

$$\frac{dF(X)}{dX} = \begin{bmatrix} \frac{\partial f_{11}}{\partial x_{11}} & \frac{\partial f_{12}}{\partial x_{12}} & \cdots & \frac{\partial f_{1n}}{\partial x_{1n}} \\ \frac{\partial f_{21}}{\partial x_{21}} & \frac{\partial f_{22}}{\partial x_{22}} & \cdots & \frac{\partial f_{2n}}{\partial x_{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{m1}}{\partial x_{m1}} & \frac{\partial f_{m2}}{\partial x_{m2}} & \cdots & \frac{\partial f_{mn}}{\partial x_{mn}} \end{bmatrix}$$

In het geval van onze functie $F(X) = X^2$, waarbij X een vierkante matrix is, is de afgeleide $\frac{dF(X)}{dX}$ gelijk aan:

$$\frac{dF(X)}{dX} = \begin{bmatrix} 2x_{11} & 2x_{12} & \cdots & 2x_{1n} \\ 2x_{21} & 2x_{22} & \cdots & 2x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 2x_{m1} & 2x_{m2} & \cdots & 2x_{mn} \end{bmatrix}$$