Resumo: Funções de Ativação em Redes Neurais

1 Função Sigmoide

Equação: $f(x) = \frac{1}{1+e^{-x}}$

- **Conceito:** Transforma entrada linear em saída não-linear entre 0 e 1
- **Aplicações:**
  - Classificação binária (probabilidades)
  - Camadas de saída em problemas de decisão sim/não
  - Detecção de fraudes bancárias
- **Lógica:** Comprime valores em intervalo (0,1), ideal para probabilidades
- **Característica:** Valores sempre positivos após passagem

2 Função Softmax

Equação: $\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{n} e^{z_k}}$

- **Conceito:** Normaliza vetor de valores para distribuição de probabilidade
- **Aplicações:**
  - Classificação multiclasse
  - Reconhecimento de imagens (múltiplas categorias)
  - Processamento de linguagem natural
- **Lógica:** Soma das saídas = 1, cada saída entre 0 e 1
- **Característica:** Considera relação entre todas as classes simultaneamente

3 Função TanH

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

- **Conceito:** Normalização entre -1 e 1
- **Aplicações:**
  - Análise de sentimentos (positivo/negativo)
  - Controle de sistemas (comandos direcionais)
  - Camadas ocultas de redes neurais
- **Lógica:** Razão entre senH e cosH, mantém simetria
- **Característica:** Saturação suave, evita mudanças bruscas

4 Função ReLU

Equação: $f(x) = max(0, x)$

- **Conceito:** Retorna 0 para valores negativos, mantém positivos
- **Aplicações:**
  - Detecção de spam em emails
  - CNNs para classificação de imagens
  - Camadas ocultas em redes profundas
- **Lógica:** Ativa neurônio apenas com entrada positiva
- **Característica:** Computacionalmente eficiente, evita vanishing gradient

5 Função LeakyReLU

Equação: $f(x) = \begin{cases} x & \text{se } x > 0 \\ \alpha x & \text{se } x \leq 0 \end{cases}$ onde $\alpha = Angulação$.

- **Conceito:** ReLU com pequena inclinação para valores negativos
- **Aplicações:**
  - Redes profundas com risco de neurônios mortos
  - Processamento de imagens com características negativas relevantes
- **Lógica:** Evita "dying ReLU"mantendo gradiente pequeno para negativos
- **Característica:** Permite backpropagation mesmo com valores negativos

6 Função ELU

Equação: $f(x) = \begin{cases} x & \text{se } x > 0 \\ \alpha(e^x - 1) & \text{se } x \leq 0 \end{cases}$

- **Conceito:** Exponencial para valores negativos
- **Aplicações:**
  - Redes muito profundas
  - Tarefas com necessidade de convergência rápida
- **Lógica:** Suaviza transição em zero, mantém média próxima de zero
- **Característica:** Reduz bias shift durante treinamento

7 Tabela Comparativa

Função	Onde Usar	Pontos Positivos	Pontos Negativos
Sigmoide	• Saída binária • Probabilidades • Gates LSTM	• Interpretação probabilística • Saída limitada (0,1) • Diferenciável	• Vanishing gradient • Computacionalmente cara • Saída não centrada em zero
Softmax	• Classificação multiclasse • Última camada • Distribuições de probabilidade	• Soma = 1 • Interpretação probabilística • Diferencia bem classes	• Computacionalmente cara • Sensível a outliers • Só para camada de saída
TanH	• Camadas ocultas • Dados normalizados • Análise sentimentos	• Centrada em zero • Saída simétrica • Convergência mais rápida	• Vanishing gradient • Computacionalmente cara • Saturação em extremos
ReLU	• Redes profundas • CNNs • Camadas ocultas	• Muito eficiente • Sem vanishing gradient • Esparsidade	• Dying ReLU • Não diferenciável em 0 • Saída não limitada
LeakyReLU	• Alternativa ao ReLU • Redes muito profundas	• Evita dying ReLU • Simples e eficiente • Sempre ativo	• Parâmetro α arbitrário • Inconsistência para negativos
ELU	• Redes muito profundas • Convergência rápida	• Suave • Robusta a ruído • Média próxima de zero	• Computacionalmente cara • Exponencial custosa