



# MATERIAL COMPLEMENTAR

Introdução ao Deep Learning

Professor Jeronimo Marcondes Pinto

# SUGESTÃO DE LEITURA

## Material disponível na internet:

- Cardon, A.; Muller, D.N. 1994. **Introdução às Redes Neurais Artificiais**. [Link](#)
  - O aluno interessado precisará fazer login com o gmail.

---

## Resumo de Conceitos

### 1. Redes Neurais Artificiais (ANNs)

São modelos computacionais inspirados na estrutura do cérebro humano, onde neurônios artificiais se conectam e processam informações. Cada neurônio recebe entradas, aplica um processamento (frequentemente através de uma função de ativação), e gera uma saída. Esse processo se repete em várias camadas, tornando possível a aprendizagem de padrões complexos a partir dos dados.

### 2. Perceptron

Um tipo básico de rede neural projetado por Frank Rosenblatt nos anos 1950, que utiliza pesos treinados para resolver problemas binários (como classificar uma entrada como “0” ou “1”). Ele é composto de um conjunto de entradas, um processamento de ponderação, uma soma e uma função de ativação. O perceptron é a base de redes mais complexas e foi o primeiro modelo prático de rede neural.

### 3. Função de Ativação

A função de ativação decide se um neurônio deve ser “ativado” ou não, ou seja, se ele deve transmitir informações para a próxima camada. Ela faz o processamento dos dados com base na combinação linear dos pesos e entradas. Exemplos: Sigmoid: Retorna um valor entre 0 e 1, ideal para saídas probabilísticas.

ReLU (Rectified Linear Unit): Retorna zero para valores negativos e o valor em si para valores positivos. É muito utilizada em redes profundas por reduzir problemas de saturação e melhorar a eficiência.

### 4. Função de Custo

Mede a diferença entre a saída esperada e a saída real da rede, avaliando o quão bom é o modelo atual. Funções de custo comuns incluem: Erro

Quadrático Médio (MSE): Utilizada em problemas de regressão para calcular o erro médio quadrado entre previsões e valores reais.

Cross-Entropy: Utilizada em problemas de classificação para medir a diferença entre a distribuição prevista e a real.

### 5. Descida do Gradiente

Técnica usada para minimizar a função de custo, ajustando os pesos da rede para reduzir o erro. Funciona calculando o gradiente da função de custo em relação a cada peso e movendo-se na direção oposta ao gradiente. A descida do gradiente estocástica (SGD) é uma variante popular que usa amostras aleatórias dos dados para atualizar os pesos, acelerando o treinamento em grandes datasets.

### 6. Backpropagation

É o processo de retropropagação do erro nas redes neurais. A partir do erro final calculado na saída, o erro é retroalimentado camada a camada até a camada inicial. Esse processo permite que cada peso seja atualizado com base em sua contribuição para o erro final, otimizando a rede como um todo.

## **7. Redes Feedforward**

Redes nas quais a informação flui em apenas uma direção, da camada de entrada para a camada de saída, sem conexões cíclicas. São mais simples e servem como base para muitos algoritmos de aprendizado profundo.

Esses conceitos formam a base de muitos algoritmos e técnicas de Deep Learning e são aplicados em áreas como reconhecimento de voz, visão computacional e processamento de linguagem natural (NLP).