

Desafio 02 - REDES NEURAIS ARTIFICIAIS (APRENDIZADO PROFUNDO)

Hanna Tatsuta Galassi

2025-10-07

Table of contents

1	Modelo MLP (Multi-Layer Perceptron)	1
1.1	Metodologia	1
1.2	Resultados da simulação	2
1.3	Classificação de novo conjunto de dígitos	2
2	Modelo CNN (Rede Neural Convolutacional)	2
2.1	Metodologia	2
2.2	Resultados da simulação	3
2.3	Classificação de novos dígitos	3
3	Discussão: vantagens e desvantagens das CNNs	3
3.1	Vantagens das CNNs	4
3.2	Desvantagens das CNNs	4

1 Modelo MLP (Multi-Layer Perceptron)

1.1 Metodologia

O treinamento do modelo MLP foi realizado utilizando o dataset MNIST. As imagens originalmente tinham tamanho 28x28 pixels, e foram pré-processadas para um vetor de 784 pixels, e normalizadas para o intervalo entre 0 e 1.

Para otimizar o processo de treinamento e evitar *overfitting*, foram implementados dois *callbacks*

- Early Stopping: o treinamento é interrompido automaticamente se a perda no conjunto de validação não apresentar melhora por 10 épocas consecutivas;
- Taxa de aprendizado adaptativa: a taxa de aprendizado é reduzida pela metade caso a perda no conjunto de validação não melhore após 5 épocas

Foram simuladas 5 topologias distintas para identificar a arquitetura com melhor performance:











1. 1 camada oculta com 512 neurônios e ativação ReLu;
2. 1 camada oculta com 32 neurônios e ativação ReLu;
3. 2 camadas ocultas com 128 e 64 neurônios, ambas com ativação ReLu;
4. 2 camadas ocultas, a primeira com 128 neurônios com ativação tanh e a segunda com 64 neurônios com ativação ReLu;
5. 3 camadas ocultas com 512, 256 e 128 neurônios, todas com ativação ReLu.

1.2 Resultados da simulação

```
--- Topologia 01 (512-relu) ---  
Acurácia de Validação Final: 0.9824  
  
--- Topologia 02 (32-relu) ---  
Acurácia de Validação Final: 0.9243  
  
--- Topologia 03 (128-relu, 64-relu) ---  
Acurácia de Validação Final: 0.9478  
  
--- Topologia 04 (128-tanh, 64-relu) ---  
Acurácia de Validação Final: 0.9440  
  
--- Topologia 05 (512-relu, 256-relu, 128-relu) ---  
Acurácia de Validação Final: 0.9719  
  
<keras.src.callbacks.history.History at 0x7022282c0520>
```

Dentre as topologias simuladas, a melhor foi a #1 (1 camada oculta, 512 neurônios, função de ativação ReLu).

1.3 Classificação de novo conjunto de dígitos

Predição: 9 Probabilidade: 0.4132 	Predição: 1 Probabilidade: 0.2582 	Predição: 2 Probabilidade: 0.3150 	Predição: 7 Probabilidade: 0.1940 
Predição: 4 Probabilidade: 0.4253 	Predição: 5 Probabilidade: 0.4794 	Predição: 5 Probabilidade: 0.3416 	Predição: 5 Probabilidade: 0.3845 
Predição: 2 Probabilidade: 0.3079 	Predição: 9 Probabilidade: 0.2380 		

O modelo obteve sucesso em apenas 3 das 10 imagens, o que indica dificuldade em generalizar para dados fora do padrão do dataset de treinamento, mesmo com alta acurácia de validação.

2 Modelo CNN (Rede Neural Convolutacional)

2.1 Metodologia

Para o modelo CNN, o pré-processamento dos dados foi diferente. As imagens foram remodeladas para um formato 4D (60000, 28, 28, 1), que preserva a estrutura espacial 2D da imagem. Os mesmos callbacks da etapa anterior foram utilizados.

Foram testadas 3 topologias CNN: 1. 2 camadas convolucionais seguidas por uma camada densa de 64 neurônios; 2. 3 camadas convolucionais e uma camada densa de 64 neurônios; 3. 2 camadas convolucionais e uma camada densa de 128 neurônios

2.2 Resultados da simulação

A performance de cada topologia CNN é apresentada a seguir:

```
--- Treinando Topologia CNN: 01 (2Conv-64Dense) ---  
Acurácia de Validação Final: 0.9917
```











```
--- Treinando Topologia CNN: 02 (3Conv-64Dense) ---  
Acurácia de Validação Final: 0.9934
```

```
--- Treinando Topologia CNN: 03 (2Conv-128Dense) ---  
Acurácia de Validação Final: 0.9860
```

A melhor topologia CNN foi a #2, composta por três camadas convolucionais. Ela apresentou uma acurácia de validação mais alta que o melhor modelo MLP.

2.3 Classificação de novos dígitos

O melhor modelo CNN foi então usado para classificar os mesmos 10 dígitos manuscritos:

Predição: 9 Probabilidade: 0.3228	Predição: 2 Probabilidade: 0.1940	Predição: 2 Probabilidade: 0.3827	Predição: 3 Probabilidade: 0.3167
			
Predição: 5 Probabilidade: 0.2307	Predição: 9 Probabilidade: 0.1896	Predição: 9 Probabilidade: 0.3251	Predição: 7 Probabilidade: 0.2821
			
Predição: 3 Probabilidade: 0.2361	Predição: 9 Probabilidade: 0.1976		
			

O modelo classificou corretamente 4 dos 10 dígitos. Embora esse resultado seja levemente melhor do que o do modelo MLP, ainda é um desempenho ruim nesse contexto de realizar a classificação de um conjunto de dados novo e com variações de estilo.

3 Discussão: vantagens e desvantagens das CNNs

As Redes Neurais Convolucionais (CNNs) são amplamente utilizadas em tarefas de visão computacional. Sua eficácia deriva de premissas sobre a natureza dos dados de imagem, o que lhes confere vantagens e desvantagens específicas quando comparadas a arquiteturas mais genéricas como o MLP.

3.1 Vantagens das CNNs

- **Eficiência de Parâmetros via Compartilhamento de Pesos:** O princípio do compartilhamento de pesos permite que um mesmo conjunto de filtros seja aplicado a diferentes regiões da imagem de entrada. Isso reduz drasticamente o número de parâmetros treináveis em comparação com uma camada densa de um MLP, diminuindo a complexidade do modelo e, conseqüentemente, a propensão ao overfitting.
- **Preservação da Estrutura Espacial e Conectividade Local:** Diferente dos MLPs, que exigem o achatamento (flattening) do dado de entrada e perdem a informação topológica, as CNNs processam as imagens em sua estrutura 2D. A conectividade local, onde cada neurônio se conecta apenas a um pequeno campo receptivo, garante a preservação das relações espaciais entre pixels, o que é fundamental para a análise de imagens.
- **Extração de Características Hierárquicas:** A composição de múltiplas camadas convolucionais permite que a rede aprenda uma hierarquia de características. As camadas iniciais tipicamente aprendem a detectar padrões de baixo nível, como bordas e texturas. As camadas subsequentes combinam essas informações para identificar padrões mais complexos e abstratos (e.g., formas, partes de objetos), criando representações de dados cada vez mais ricas.

3.2 Desvantagens das CNNs

- **Elevado Custo Computacional:** Embora eficientes em parâmetros, as operações de convolução são computacionalmente intensivas. O treinamento de CNNs profundas é significativamente mais lento que o de MLPs, o que dificulta sua utilização quando não se tem acesso a computadores com especificações robustas.
- **Complexidade no Design da Arquitetura:** A performance de uma CNN é altamente sensível a um grande conjunto de hiperparâmetros, incluindo a profundidade da rede, o tamanho do kernel, o passo (stride), o preenchimento (padding) e o tipo de camada de agrupamento (pooling). O projeto de uma topologia ótima é um processo empírico, não trivial, que exige extensiva experimentação.