

Universidade Estadual de Feira de Santana Engenharia de Computação EXA868 Inteligência Artificial Não Simbólica Prof. Matheus Giovanni Pires



## **EPC 3**

## Data de Entrega: 31/03/2023

Os seres humanos são ótimos no reconhecimento de padrões e são capazes de fazer isso quase que imediatamente e praticamente sem esforço. Nós podemos, por exemplo, reconhecer um rosto familiar de uma pessoa muito embora esta pessoa tenha envelhecido desde o último encontro, identificar uma pessoa familiar pela sua voz ao telefone, apesar de uma conexão ruim e distinguir um ovo fervido que é bom de um ruim pelo seu cheiro. Os humanos realizam o reconhecimento de padrões através de um processo de aprendizagem, e assim acontece com as Redes Neurais Artificiais [1].

O reconhecimento de padrões é formalmente definido como o processo pelo qual um padrão/sinal recebido é atribuído a uma classe dentre um número predeterminado de classes. Uma rede neural realiza o reconhecimento de padrões passando inicialmente por uma seção de treinamento, durante a qual se apresenta repetidamente à rede um conjunto de padrões de entrada junto com a categoria à qual cada padrão particular pertence. Mais tarde, apresenta-se à rede um novo padrão que não foi visto antes, mas que pertence à mesma população de padrões utilizada para testar a rede. A rede é capaz de identificar a classe daquele padrão particular por causa da informação que ela extraiu dos dados de treinamento [1].

Neste EPC, você irá projetar Redes Perceptron Multicamadas (PMC) para a resolução do problema de classificação *Iris Plants* (dataset disponível no repositório *Keel Dataset* https://sci2s.ugr.es/keel/datasets.php).

1. Treine e valide a melhor topologia usando 10-folds cross-validation usando o algoritmo de aprendizagem backpropagation padrão, inicializando as matrizes de pesos com valores aleatórios entre 0 e 1. Utilize a função de ativação logística (sigmoid) para todos os neurônios, taxa de aprendizado  $\eta = 0.1$ , precisão  $\varepsilon = 10^{-6}$ e  $\beta = 0.5$ . Para cada topologia avaliada, registre em uma tabela a média e o desvio padrão das seguintes medidas: EQM, Número de épocas e Acurácia (porcentagem de acerto do conjunto de validação).

| Topologia EQM Épocas Acurácia |         |                |        |
|-------------------------------|---------|----------------|--------|
| 4-5-3                         | 0.02829 | 961 épocas     | 0.9333 |
| 4-10-3                        | 0.03388 | 364 épocas     | 1.0    |
| 4-15-3                        | 0.0780  | 5935<br>épocas | 1.0    |
| 4-20-3                        | 0.0398  | 427 épocas     | 1.0    |

Função de ativação logística:

$$1 + \diamondsuit \diamondsuit \diamondsuit \diamondsuit \diamondsuit \diamondsuit \diamondsuit \diamondsuit (\diamondsuit \diamondsuit) = \diamondsuit \diamondsuit . \diamondsuit \diamondsuit (\diamondsuit \diamondsuit) . (1 - \diamondsuit \diamondsuit (\diamondsuit \diamondsuit))$$

Dado que o problema se configura como um típico processo de classificação de padrões, implemente uma rotina que faz o pós-processamento das saídas fornecidas pela rede (números reais) para números inteiros. Utilize o critério do arredondamento simétrico:

Página 1 de 2

- 2. Repita o experimento do item 1 utilizando o algoritmo de aprendizagem *backpropagation* com momentum. Fator de momentum  $\alpha = 0.9$
- 3. Após encontrar a melhor topologia, uma usando o *backpropagation padrão* e outra usando *backpropagation com momentum*, teste as redes usando os conjuntos de teste. Registre em uma tabela a média e o desvio padrão das taxas de acerto (%).

| Algoritmo de treinamento Topologia Acurácia |       |              |  |  |
|---|-------|--------------|--|--|
| Backpropagation padrão                      | 4-5-3 | (89.3% ± 8%) |  |  |
| Backpropagation com momentum                | 4-5-3 | (90.6% ± 8%) |  |  |

4. Explique o que são situações de *underfitting* e *overfitting*, descrevendo-se também os meios para as suas detecções. Quais são as possíveis soluções para cada situação?

Underfitting: Incapacidade do modelo de representar o fenômeno estudado. Pode ser identificado quando o desempenho dos modelos é muito baixo. Para tentar solucionar isso pode-se avaliar configurações paramétricas que possibilitem o aumento da capacidade de representação do modelo, como no nosso caso, através do aumento do número de camadas ou de neurônios por camada.

Overfitting: Uma situação de super-adequação aos dados, quando o modelo aprende a memorizar os dados aos quais ele é apresentado. É frequentemente associada à baixa disposição de exemplos de um conjunto de dados ou exemplos que não sejam representativos o suficiente para possibilitar o aprendizado e pode ser identificado quando há uma divergência de desempenho muito acentuada entre o treinamento e o teste, isto é, quando o desempenho de treinamento é muito superior ao do teste. Outra maneira de verificar isto é através da realização de validação cruzada. Através dessa abordagem, uma alta variação de desempenho entre os folds pode dar um indicativo de que o modelo esteja sofrendo com overfitting. Para contornar isso pode-se tentar, se possível, realizar mais coletas de dados, com dados mais representativos e/ou utilizar aumento de dados também.

## **OBSERVAÇÕES**

- 1. O EPC deve ser realizado individualmente.
- 2. Os resultados devem ser entregues em sequência, ou seja, de acordo com a numeração do

## REFERÊNCIAS

[1] S. O. Rezende, Sistemas Inteligentes: fundamentos e aplicações, Manole, Barueri-SP, 2003.