ICMC-USP

Lista de Exercícios - Parte 1

SCC-0570 - Redes Neurais

1o. Semestre de 2019 - Prof. João Luís - PAE: Fernando Aguiar

1. Um neurônio tem uma função de ativação $\varphi(v)$ definida pela função logística da Eq. (1), onde v é o campo local induzido, e o parâmetro de inclinação a está disponível para ajuste. Seja $x_1, x_2, ..., x_m$ os sinais de entrada aplicados aos nós fonte do neurônio, e b o *bias*. Para conveniência de apresentação, deseja-se absorver o parâmetro de inclinação a no campo local induzido v escrevendo a Eq. (2).

$$\varphi(v) = \frac{1}{(1 + exp(-av))} \tag{1}$$

$$\varphi(v) = \frac{1}{(1 + exp(-v))} \tag{2}$$

   Como você modificaria as entradas $x_1, x_2, ..., x_m$ para a Eq. (2) produzir a mesma saída obtida pela Eq. (1)? Justifique sua resposta.

2. Um neurônio j recebe entradas de outros 4 neurônios cujos níveis de atividades são: 10, -20, 4 e -2. Os pesos sinápticos respectivos são: 0.8, 0.2, -1.0 e -0.9. Calcule a saída do neurônio j para as seguintes situações:
   (a) o neurônio é linear, isto é, sua função de transferência ou ativação é linear.
   (b) o neurônio é representado por um modelo de McCulloch-Pitts. Assuma que o *bias* aplicado ao neurônio inicialmente é zero e depois é igual a 0.5.
   (c) o neurônio é baseado na função logística da equação 2.

3. Uma rede *feedforward* totalmente conectada tem 10 nós fonte, 2 camadas escondidas, uma com 4 neurônios e a outra com 3 neurônios, e um único neurônio de saída. Construa um grafo arquitetural desta rede.

4. Considere uma rede *feedforward* multicamadas, com todos os neurônios operando em suas regiões lineares (segmento linear de suas funções de ativação não-lineares, e.g. sigmóides). Justifique a afirmação de que tal rede é equivalente a uma rede *feedforward* de única camada.

5. Desenhe o grafo arquitetural de uma rede recorrente totalmente conectada com 5 neurônios, mas nenhuma auto-retro-alimentação.

6. Uma rede recorrente tem 3 nós fonte, 2 neurônios escondidos e 4 neurônios de saída. Construa um grafo arquitetural que descreva tal rede.

 Continua na próxima página...

7. Como se dá o aprendizado nas redes neurais artificiais?

8. Explique, em palavras, o que representa a regra delta.

9. O aprendizado Hebbiano é considerado aprendizado não supervisionado, apesar do treinamento ser feito com pares entrada-saída. Comente esta afirmação [1].

10. Quais são os principais paradigmas de aprendizado? Descreva características de cada um deles e indique a quais tipos de problemas cada um se aplica [1].

11. Diferencie aprendizado não supervisionado de aprendizado supervisionado [1].

12. Explique a diferença entre a hipótese de Hebb

$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n) \tag{3}$$

e a hipótese da covariância

$$\Delta w_{kj} = \eta (x_j - \bar{x})(y_k - \bar{y}) \tag{4}$$

13. Como você classificaria o aprendizado competitivo? Supervisionado? Não supervisionado? Outro?

14. O aprendizado por reforço pode ser visto como um tipo de aprendizado supervisionado ao se utilizar o sinal de reforço como sendo alguma medida da distância entre a saída atual do sistema e a saída desejada. Discuta a relação entre aprendizado por reforço e aprendizado supervisionado [1].

15. O perceptron é um classificador de padrões linear. Justifque esta afirmação.

16. O perceptron pode ser usado para executar numerosas funções lógicas. Demonstrar a implementação das funções lógicas binárias AND, OR e NOT. Entretanto, uma limitação básica do perceptron é que ele não pode implementar o OU-EXCLUSIVO. Explique a razão para esta limitação.

17. Considere um conjunto de pontos referentes a uma classe $\mathscr{C}_1$ está distribuído próximo ao ponto (8,5) no plano XY. Os pontos não pertencentes à classe $\mathscr{C}_1$ são externos à região em que os pontos de $\mathscr{C}_1$ são amostrados e correspondem à classe $\mathscr{C}_2$. Mostre a solução de uma rede perceptron para fazer a separação das duas classes $\mathscr{C}_1$ e $\mathscr{C}_2$, sem utilizar algoritmos de treinamento convencionais. Obtenha a solução fazendo apenas a interpretação geométrica do problema. Discuta também como a variância da distribuição dos elementos da classe $\mathscr{C}_1$ afeta o problema da classificação.

18. Verifique que as equações 5, 6, 7 e 8 abaixo

$$sgn(v) = \begin{cases} +1, & se\ v > 0 \\ -1, & se\ v < 0 \end{cases} \tag{5}$$

 Continua na próxima página...

$$y(n) = sgn[\mathbf{w}^T(n)\mathbf{x}(n)] \tag{6}$$

$$d(n) = \begin{cases} +1, & se\ \mathbf{x}(n)\ pertence\ a\ classe\ \mathscr{C}_1 \\ -1, & se\ \mathbf{x}(n)\ pertence\ a\ classe\ \mathscr{C}_2 \end{cases} \tag{7}$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta[d(n) - y(n)]\mathbf{x}(n) \tag{8}$$

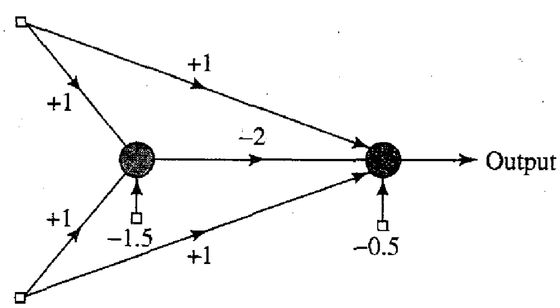
que resumem o algoritmo de convergência do perceptron, são consistentes com:

1. "Se o n-ésimo termo do conjunto de treinamento $\mathbf{x}(n)$ é corretamente classificado por $\mathbf{w}(n)$ computado na n-ésima iteração do algoritmo, nenhuma correção é feita ao vetor de pesos de acordo com a regra
   (a) $\mathbf{w}(n+1) = \mathbf{w}(n)$, se $\mathbf{w}^T\mathbf{x}(n) > 0$ e $\mathbf{x}(n)$ pertence à classe $\mathscr{C}_1$.
   (b) $\mathbf{w}(n+1) = \mathbf{w}(n)$, se $\mathbf{w}^T\mathbf{x}(n) \leq 0$ e $\mathbf{x}(n)$ pertence à classe $\mathscr{C}_2$.
2. Caso contrário, o vetor de pesos é atualizado de acordo com a regra
   (a) $\mathbf{w}(n+1) = \mathbf{w}(n) - \eta(n)\mathbf{x}(n)$, se $\mathbf{w}^T(n)\mathbf{x}(n) > 0$ e $\mathbf{x}(n)$ pertence à classe $\mathscr{C}_2$.
   (b) $\mathbf{w}(n+1) = \mathbf{w}(n) + \eta(n)\mathbf{x}(n)$, se $\mathbf{w}^T(n)\mathbf{x}(n) \leq 0$ e $\mathbf{x}(n)$ pertence à classe $\mathscr{C}_1$.

   onde $\eta(n)$ controla o ajuste aplicado ao vetor de pesos na iteração n."

19. A figura 1 abaixo mostra uma rede neural com um único neurônio escondido. Mostre que essa rede resolve o problema do XOR construindo (a) regiões de decisão e (b) uma tabela-verdade para a rede.

Figura 1: Rede Neural com único neurônio escondido.



20. Por que o *back-propagation* é chamada de regra delta generalizada?

21. O que é *overfitting*?

22. [EXPERIMENTO COMPUTACIONAL] Investigue o emprego do aprendizado *back-propagation* usando uma não-linearidade sigmoidal para realizar mapeamentos um-a-um, como descritos abaixo:

(1)	$f(x) = \frac{1}{x}$	$1 \leq x \leq 100$
(2)	$f(x) = log_{10}x$	$1 \leq x \leq 10$
(3)	$f(x) = exp(-x)$	$1 \leq x \leq 10$
(4)	$f(x) = sen\ x$	$0 \leq x \leq \frac{\pi}{2}$

Para cada mapeamento, faça o seguinte:

a) Estabeleça dois conjuntos de dados, um para o treinamento da rede e outro para teste.

b) Use o conjunto de dados de treinamento para computar os pesos sinápticos da rede, assumindo haver uma única camada escondida.

c) Avalie a acurácia computacional da rede usando os dados de teste.

Use apenas uma camada escondida mas com número variável de neurônios escondidos. Investigue como a performance da rede é afetada variando-se o tamanho da camada escondida.

Referências

[1] A. P. Braga, A. P. L. F. Carvalho, T. B. Ludermir, *Redes Neurais Artificiais - Teoria e Aplicações*, 2a. edição. LTC, 2007.

[2] S. Haykin, *Neural networks - a comprehensive foundation*, 2nd. ed. Prentice Hall, 1999.

[3] R. A. F. Romero, "SCC 5809 - Redes Neurais," Slides, 2o. semestre de 2010.

 Final da Lista.