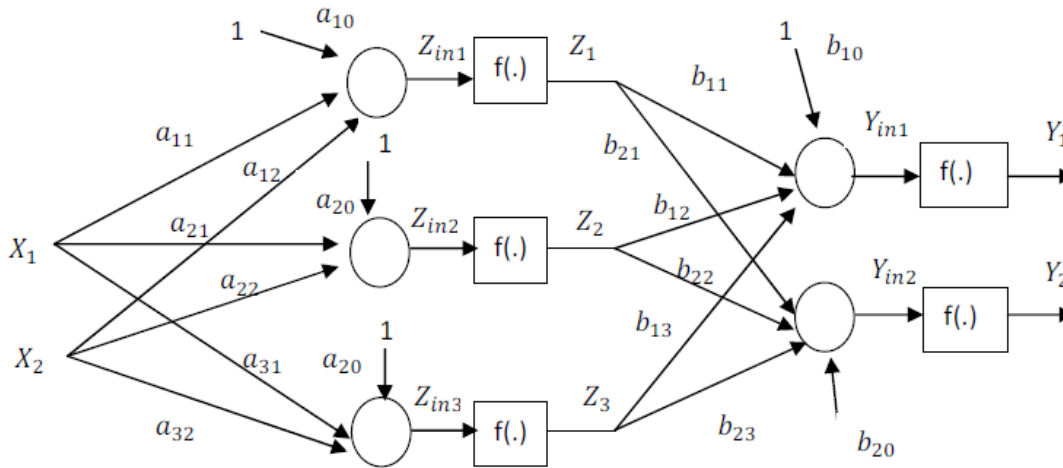


1ª Questão (1.5 pontos) A figura abaixo mostra uma rede neural feedforward de múltiplas camadas. Suponha que sua taxa de aprendizado seja 0,5 e a função de ativação é a função logística. Os pesos iniciais e bias são dados logo abaixo na sequência da figura



$$A = \begin{bmatrix} a_{11} & a_{12} & a_{10} \\ a_{21} & a_{22} & a_{20} \\ a_{31} & a_{32} & a_{30} \end{bmatrix} = \begin{bmatrix} 0.1 & -0.1 & -0.1 \\ 0.1 & 0.1 & -0.1 \\ -0.1 & -0.1 & 0.1 \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{10} \\ b_{21} & b_{22} & b_{23} & b_{20} \end{bmatrix} = \begin{bmatrix} 0.1 & 0.0 & 0.1 & -0.1 \\ -0.1 & 0.1 & -0.1 & 0.1 \end{bmatrix}$$

Assumindo que a entrada para o treinamento seja $X = (0,1)$, cuja classificação esperada é $(1, 0)$, pede-se

- (0.5 ponto) Calcule a saída da rede neural e o erro em cada neurônio de saída.
- (1.0 ponto) Mostre os cálculos que o algoritmo backpropagation executará na primeira época para os pesos $b_{11}, b_{22}, b_{23}, a_{11}, a_{22}$

2ª Questão (1,0 Ponto) Assinale verdadeiro (V) ou falso (F). Lembre-se que um item assinalado incorretamente anula um item corretamente

(.....) A escolha adequada da taxa de aprendizado em Redes Neurais Artificiais é muito importante para assegurar a estabilidade da convergência do processo de aprendizado iterativo, pois taxas altas permitem um aprendizado mais lento porém mais consistente, mas com o perigo de cair em mínimos locais, enquanto taxas baixas permitem um aprendizado mais rápido a custo muitas vezes de desestabilização (oscilação)

(.....) A ativação de um neurônio na rede MLP se dá pelo produto interno entre seu vetor de pesos e o vetor de entradas, seguida pela aplicação da função de ativação, geralmente do tipo sigmoidal. Por outro lado, a ativação de um neurônio na rede RBF se dá pelo cálculo da distância euclidiana entre o vetor de pesos do neurônio e o vetor de entradas. Quanto mais distante o vetor de entrada do vetor de pesos, menor a ativação do neurônio.

(.....) as redes MLP tendem a se dar melhor no caso de número elevado de entradas, quando comparado às redes RBF. Conforme aumenta o número de entradas, o número de funções de base radial tende a crescer exponencialmente, caso se queira manter o mesmo nível de desempenho. Essa lei é conhecida como “maldição da dimensionalidade”.

(.....) motivado pelas decisões de projeto, o treinamento de uma rede neural RBF se dá em um único passo de cálculo (usando uma fórmula matemática fechada), representado pela pseudo-inversão de uma matriz e por produtos entre matrizes e vetores. Com isso, o projetista da rede RBF deve definir o número de neurônios, o centro das funções de base radial e parâmetros de dispersão desta função de base radial. Por outro lado, o treinamento de uma rede neural MLP envolve aplicações iterativas de um processo de ajuste incremental do vetor de pesos, sendo necessário definir a cada iteração um passo e uma direção de ajuste.

() Holdout reserva uma certa quantidade de dados para treinamento e o restante para teste (podendo ainda usar parte para validação). Comumente esta estratégia usa 1/3 dos dados para teste e o restante para treinamento, escolhido

randomicamente. Para conjunto de dados randômicos é interessante assegurar que a amostragem randômica seja feita de tal maneira que garanta que cada classe é apropriadamente representada tanto no conjunto de treinamento quanto no conjunto de teste. Este procedimento é chamado de holdout repetitivo. Se for realizado apenas uma divisão do conjunto de dados, a estimativa da taxa de erro vai ser enganosa se acontecer de termos uma divisão ruim. Visando amenizar tendências, emprega-se holdout estratificado, o qual consiste em repetir todo o processo de treino e teste várias vezes com diferentes amostragens randômicas.

() O bootstrap é em um procedimento estatístico de amostragem com reposição. Considerando um conjunto de dados com n instâncias, n instâncias são escolhidas aleatoriamente. Uma instância não é retirada do conjunto de dados original quando ela é escolhida para compor o conjunto de treinamento, ou seja, a mesma instância pode ser selecionada várias vezes durante o procedimento de amostragem.

() No n fold cross validation o conjunto de dado é dividido em n partições de tamanhos aproximadamente iguais e, de maneira rotativa, cada uma delas é usada para teste enquanto as restantes são usadas para treinamento. Este procedimento é repetido n vezes. Para conjunto de dados pequeno, geralmente n é escolhido igual ao número de instâncias no conjunto de dados. Este é também conhecido como leave one out.

(.....) Quando o objetivo do SOM é a visualização do conjunto de dados pode-se desconsiderar a necessidade de quantização do espaço, mas não a necessidade de redução de dimensão.

(.....) K-prototypes visar neutralizar a desvantagem do Kmeans no que diz respeito a sua sensibilidade a ruídos

(.....) No **Ressubstituion** o conjunto de treinamento empregado para treinar o classificador é também usado para testar a sua performance. Os desempenhos computados com este método são otimistas e tem grande bias . Uma vez que não há divisão do conjunto de dados, pode ser para conjuntos de dados pequenos e grandes e para distribuição balanceada ou não.

3ª Questão) (2,0 Pontos) Considere o problema de aprender o conceito sobre pacientes doentes (+) ou sadios (-). Foram coletados os seguintes exemplos

- (1,0 ponto) Utilizando o algoritmo de indução de árvores de decisão, construa a árvore correspondente (sem poda e sem número mínimo de exemplos em cada folha), utilizando o critério de ganho de informação para selecionar atributos para este conjunto de exemplos. Anote, para cada nível da árvore de decisão, o valor do ganho de informação calculado para cada atributo, bem como aquele escolhido para particionar os exemplos. Se houver empate entre valores do ganho de informação, escolha o primeiro (na ordem da tabela acima).
- (0,5 ponto) Análise de desempenho: use a árvore de decisão produzida em (a) para classificar os exemplos do conjunto de teste. Informe a precisão e matriz de confusão da árvore para esses exemplos. Discuta resumidamente os resultados.
- (0,5 ponto) Aplique o algoritmo KNN ($k = 2$) ao problema acima e monte a matriz de confusão para os valores de k . Faça a codificação que julgar necessário.

Usando distância euclidiana $dist(i, j) = ExameA_i \neq ExameA_j + ExameB_i \neq ExameB_j$

Conjunto de Treinamento			
ID	Exame A	Exame B	Classe
1	H	C	+
2	R	B	+
3	R	C	+
4	H	B	+
5	J	B	+
6	R	C	-
7	J	C	-
8	J	B	-
9	H	C	+
10	J	C	-
11	R	C	-
12	J	B	+
13	R	C	-

Conjunto de Teste			
14	R	B	+
15	J	B	-
16	J	C	-
17	R	C	+
18	H	C	+

4ª Questão) (2,5 Pontos) Abaixo segue o arquivo lar_lim.csv que contém o padrão de consumo per capita mensal de laranja e limão das cidades A, B, C, D, E, F, G. O arquivo serve de entrada para os algoritmos de agrupamento

Dados

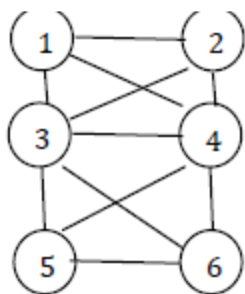
Cidade	Laranja	Limão
A	55	43
B	80	63
C	85	50
D	58	38
E	82	55
F	66	42

Matriz de distância

	A	B	C	D	E	F
A	0	32.0	30.8	5.8	29.5	11.0
B	32.0	0	13.9	33.3	8.2	25.2
C	30.8	13.9	0	29.5	5.8	20.6
D	5.8	33.3	29.5	0	29.4	8.9
E	29.5	8.2	5.8	29.4	0	20.6
F	11.0	25.2	20.6	8.9	20.6	0

- (1,0 Ponto) Aplique a clusterização hierárquica (single link) e apresente o dendograma gerado.
- (0,75 Ponto) Aplique Kmeans para K = 2 (Centróide Iniciais A , B). Apresente uma iteração
- (0,75 Ponto) Considerando os pontos A e B, como medóides iniciais, verifique se há algum ponto que pode substituir o medóide A.

5ª Questão) (2,0 Ponto) Para o conjunto de dados descrito na questão anterior (lar_lim.csv), uma rede neural SOM deve ser treinada. Foi gerado um grid bi-dimensional e os pesos iniciais para cada neurônio é dado na tabela abaixo.



Neurônio	Coordenada X	Coordenada Y
1	50	35
2	70	35
3	85	35
4	50	50
5	70	50
6	85	50

Considerando $h(w_i, w_j) = \exp\left(\frac{-\sum_k (w_{ik} - w_{jk})^2}{2\sigma^2}\right)$, e a atualização dos pesos dada pela seguinte regra

$$w_i(\text{new}) = w_i(\text{old}) + \alpha(t)h(w_i, w_j) [x_i - w_i(\text{old})],$$

onde i é o neurônio vencedor e $\alpha(0) = 1$. Pede-se

- (1,0 Ponto) Para apenas as duas primeiras instâncias, calcule o neurônio vencedor e atualize seus pesos e dos neurônios, considerando vizinhança 1.
- (1,0 Ponto) **Com base nos pesos da letra a)**, calcule a matriz U.

6ª Questão) (1.0 pontos) Considere o problema do Ou-Exclusivo, definido em sala de aula. Suponha que uma rede neural do tipo RBF, com dois neurônios na camada interna, tenha sido empregada para este problema

- (0.5 ponto) Apresente a representação gráfica da rede para este problema, definindo o número de entradas e saída.
- (0.5 ponto) Considere $h_1(x)$, $h_2(x)$ descrito abaixo, como a função de ativação dos neurônios 1 e 2 da camada escondida da rede RBF. Supondo que o vetor de pesos da camada de saída seja $w = [w_0 \ w_1 \ w_2] = [0.1 \ 0.4 \ 0.3]$. Calcule a saída da rede para o problema do Ou-Exclusivo

$$h_1(x) = \exp\left(-\frac{(x_1 - c_{11})^2}{2\sigma_1^2} - \frac{(x_2 - c_{12})^2}{2\sigma_1^2}\right), h_2(x) = \exp\left(-\frac{(x_1 - c_{21})^2}{2\sigma_2^2} - \frac{(x_2 - c_{22})^2}{2\sigma_2^2}\right)$$

$$\text{onde } c = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ e } \sigma_1 = 1, \sigma_2 = 1$$