



# Redes Neurais Artificiais

## AULA 12 – Redes ART – Arquitetura e Aspectos de Treinamento –

Prof. Ivan Nunes da Silva



### 1. Introdução

#### *Aspectos introdutórios da arquitetura*

- A Teoria da Ressonância Adaptativa (*Adaptive Resonance Theory* // ART) foi desenvolvida a partir da observação de alguns fenômenos biológicos envolvendo a visão, a fala, o desenvolvimento cortical e as interações cognitivo-emocionais.
- A motivação desta teoria é baseada em três princípios biológicos que são destacados pelas seguintes características:
  - **Normalização de Sinais** → Capacidade dos sistemas biológicos em se adaptar em ambientes que se alteram a todo momento. Como exemplo, tem-se o sistema oftalmológico humano em que se adapta rapidamente frente à quantidade de luz disponível.
  - **Intensificação de Contraste** → Potencialidade de identificar detalhes sutis a partir de sucessivas observações realizadas sobre o ambiente. Como exemplo, o sistema respiratório é capaz de diferenciar de forma quase que instantânea um ambiente limpo que começou a ser poluído com monóxido de carbono.
  - **Memória de Curto Prazo** → Possibilidade de armazenar momentaneamente as informações sensoriais advindas do mecanismo de intensificação de contraste, antes que as mesmas possam ser decodificadas visando à tomada de decisões.

# 1. Introdução

## Principais atributos da rede ART

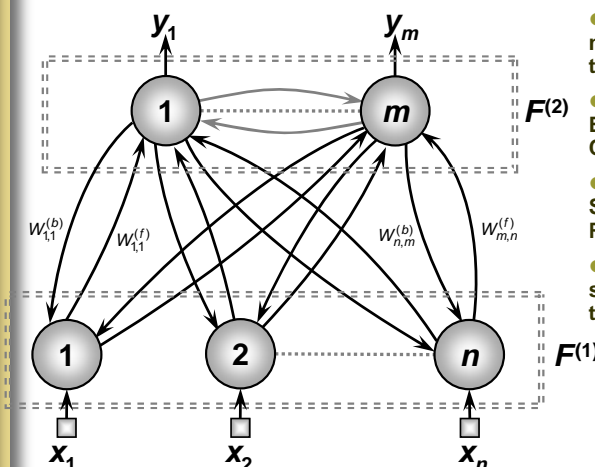
- Os principais atributos das redes ART é a sua habilidade de aprender novos padrões sem destruir os conhecimentos já anteriormente extraídos, por ocasião da apresentação de outras amostras.
- Esta característica está diretamente associada ao intrigante dilema plasticidade/estabilidade:
  - O sistema deve ser flexível (adaptativo) o suficiente para incorporar mudanças ocorridas no ambiente.
  - O sistema deve ser ainda bem estável a fim de preservar o conhecimento já adquirido no decorrer do tempo
- Principais topologias das redes ART:
  - **ART-1** → Padrões c/ entradas binárias e treinamento não-supervisionado.
  - **ART-2** → Padrões com entradas binárias ou analógicas (contínuas) e treinamento não-supervisionado.
  - **ART-3** → Padrões com entradas binárias ou analógicas e treinamento não-supervisionado, usando-se aqui topologia multinível e “transmissores químicos” para controle do processo de busca pela melhor solução.
  - **ART-Map** → Padrões com entradas binárias ou analógicas e treinamento supervisionado (tempo-real), usando-se 2 redes ART em sua estrutura.
  - **Fuzzy-ART** → Versão fuzzy da ART-MAP e que possui também aprendizado supervisionado

3

# 2. Redes ART-1

## Estrutura topológica

- Em função da grande aplicabilidade, abordar-se-á nesta aula somente a topologia ART-1, que são aplicadas em classificação de padrões.
- As amostras (vetores) de entrada são compostas por valores binários.



- Possui 2 camadas de neurônios  $\{F^{(1)} \text{ e } F^{(2)}\}$ , totalmente conectadas.

- $F^{(1)}$  → Camada de Entrada ou Camada de Comparação.

- $F^{(2)}$  → Camada de Saída ou Camada de Reconhecimento.

- O aprendizado é não-supervisionado (não se tem a saída desejada).

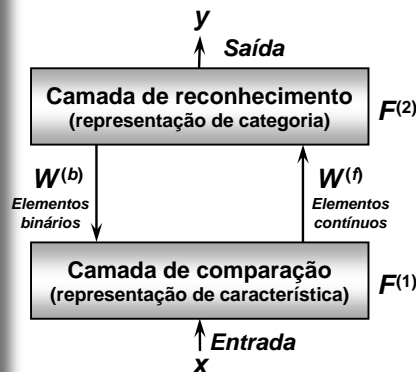
4

## 2. Redes ART-1

### Estrutura topológica (camada de comparação)

#### Camada de Comparação $\{F^{(1)}\}$

- Seja a apresentação de uma amostra representada pelo vetor  $\{x\}$ .



#### Aspectos da Camada $F^{(1)}$ :

- A camada  $F^{(1)}$  possui uma quantidade de neurônios igual à dimensão das entradas (vetor  $x$ ).
- A finalidade de  $F^{(1)}$  é também de repassar tais padrões originais para os neurônios da camada  $F^{(2)}$ .
- As informações realimentadas (*feedback*) são os sinais de entrada de  $F^{(1)}$ , sendo ponderados pela sua matriz de pesos  $\{W^{(b)}\}$ , composta por elementos binários.
- A missão da camada  $F^{(1)}$  consiste então de realizar uma operação de comparação a fim de verificar se o vetor de entrada  $\{x\}$ , previamente associado a uma das classes representadas pelos neurônios da camada de saída, possui similaridade com o conjunto de vetores já armazenados na mesma.

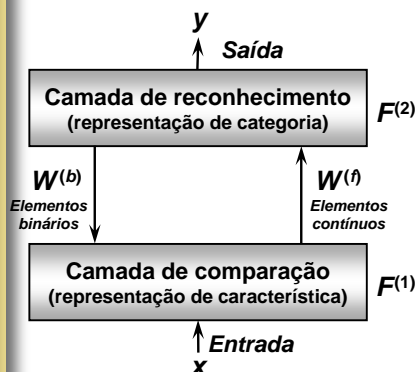
5

## 2. Redes ART-1

### Estrutura topológica (camada de reconhecimento)

#### Camada de Reconhecimento $\{F^{(2)}\}$

- A missão de  $F^{(2)}$  é reconhecer/incorporar o conhecimento trazido pela nova entrada p/ uma das classes representadas por seus neurônios.
- Portanto, o número de neurônios desta camada é dinâmica, variando-se de acordo com a necessidade de se adicionar novos neurônios, com a finalidade de que novas classes possam ser representadas.
- Esta característica é um dos principais atrativos de redes ART.



#### Aspectos da Camada $F^{(2)}$ :

- Os neurônios da camada  $F^{(2)}$  são competitivos entre si.
- O fluxo de informações de  $\{F^{(2)}\}$  realimenta os neurônios de  $\{F^{(1)}\}$ .
- Sua matriz de pesos  $\{W^{(f)}\}$ , formada de elementos reais, fica incumbida de ponderar as características das componentes do vetor de entrada  $\{x\}$ :
  - O objetivo é estimular que um dos neurônios representativos das classes possa ser então ativado como sinal de reconhecimento do referido vetor.
  - Similar à rede de Kohonen, haverá apenas um vencedor em cada ciclo.

6

## 2. Redes ART-1

### Princípio da Ressonância Adaptativa (I)

- O funcionamento da rede ART-1 é baseado no próprio princípio da ressonância adaptativa, quando da apresentação de um novo vetor (representando uma amostra) em sua camada de entrada.
- Em primeira instância, cada vetor  $\{x\}$  apresentado será propagado de  $F^1$  para  $F^2$ , sendo suas componentes ponderadas pela respectiva matriz de pesos  $\{W^0\}$  da camada  $F^2$ .
  - O intuito aqui é extrair as suas características discriminantes, as quais visam estimular uma das classes representadas pelos seus neurônios.
- Em segunda instância, os neurônios de  $F^2$ , os quais estarão representando as classes do problema a ser mapeado, entrarão, portanto, em competição.
  - O objetivo aqui é verificar qual deles possui maior proximidade com aquele vetor que acabou de chegar, declarando-se um vencedor p/ a competição.
- Em terceira instância, o neurônio ganhador encaminha a classe vencedora de volta à camada de entrada  $F^1$ :
  - Durante este processo, os neurônios da camada  $F^1$ , por intermédio dos pesos armazenados na matriz  $W^0$ , realizam então o teste de similaridade a fim de verificar a aderência daquele vetor de entrada à classe representada pelo neurônio vencedor de  $F^2$ .

7

## 2. Redes ART-1

### Princípio da Ressonância Adaptativa (II)

- Em quarta instância, em decorrência do teste de similaridade, se o vetor de pesos do vencedor for “parecido” com o respectivo vetor de entrada, levando-se em conta um limiar de aderência (parâmetro de vigilância), haverá então o processo de RESSONÂNCIA entre ambos.
  - Em consequência, o referido vetor de entrada será associado àquela classe em que o vencedor esteja representando.
  - A finalização desse ciclo de reconhecimento conduz à atualização de ambas as matrizes de pesos.
- Caso contrário, em quinta instância, o atual vencedor (ativo), embora tenha ganhado a competição em  $F^2$ , não é aquele mais apropriado para representar a classe associada ao referido vetor de entrada.
  - Nesta situação, no próximo ciclo, o mesmo deve ser então desabilitado para que outro neurônio da camada  $F^2$  possa então ter a chance de vencer a competição, repassando-se (em seguida) a nova classe vencedora para a realização do teste de similaridade.
  - Tal processo se repete a cada novo ciclo até que haja definição de uma classe que seja compatível com o referido vetor de entrada.

8

## 2. Redes ART-1

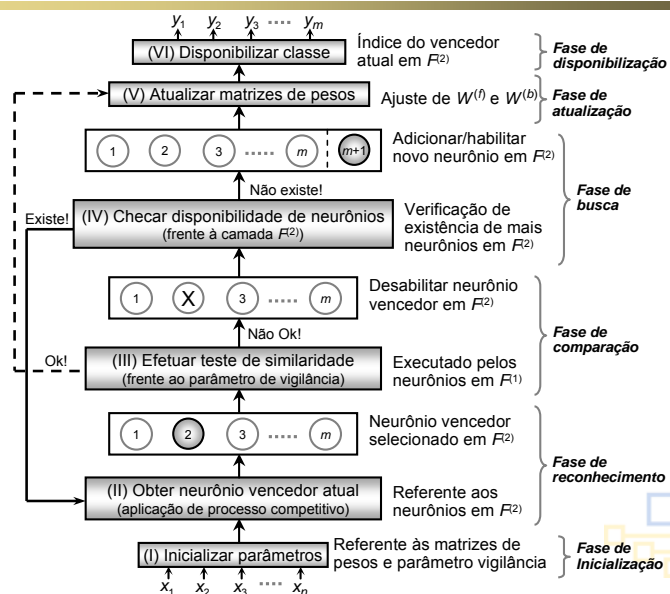
### Princípio da Ressonância Adaptativa (III)

- Finalmente, em última instância, caso nenhum dos neurônios de  $F^2$  seja aprovado no teste de similaridade, um novo neurônio deve ser então adicionado/habilitado na camada  $F^2$  a fim de identificar esta outra classe representada pelo vetor de entrada em questão.
  - De fato, tratar-se-á então de uma nova classe, pois todas as outras classes já existentes, as quais estão sendo representadas pelos outros neurônios de  $F^2$ , não são passíveis de englobar as novas características de conhecimento/ comportamento trazido pelo referido vetor de entrada.
- Na realidade, o processo de ressonância realiza basicamente um teste de hipótese.
  - Nesta circunstância, uma hipótese sobre uma classe apropriada para incorporar o vetor de entrada é formulada pela camada de reconhecimento  $F^2$ .
  - Em seguida, a camada de comparação  $F^1$  entra em operação a fim de validar a hipótese proposta, ou seja, se o vetor de entrada tiver um grau de similaridade aceitável à classe sugerida, há então a validação da hipótese;
    - Caso contrário, a hipótese previamente formulada estava equivocada e outras classes (neurônios) devem ser então avaliadas, ou até mesmo criadas, nos próximos ciclos.

9

## 3. Aprendizado da ART-1

### Fluxograma de funcionamento



10

### 3. Aprendizizado da ART-1

#### Etapa (I) – Inicialização de Parâmetros

- Esta etapa consiste de atribuir valores iniciais p/  $W^{(f)}$ ,  $W^{(b)}$  e  $\{\rho\}$ .
  - A missão do parâmetro de vigilância é arbitrar o teste de similaridade a ser executado pelos neurônios da camada de entrada  $F^{(1)}$ .
- A matriz  $W^{(f)}$ , referente à saída  $F^{(2)}$ , deve ser inicializada com valores pequenos a fim de evitar saturação das respostas. Para tanto, tem-se:

$$W_{ji}^{(f)} = \frac{1}{1+n}, \text{ com } \begin{cases} j = 1, \dots, m \\ i = 1, \dots, n \end{cases}$$

– onde  $n$  é o número de neurônios em  $F^{(1)}$  e  $m$  é a quantidade de neurônios habilitados em  $F^{(2)}$ .

- A matriz  $W^{(b)}$ , referente a  $F^{(1)}$ , é inicializada com valores unitários, que indica que cada neurônio pertencente à camada  $F^{(2)}$  está a princípio habilitado para enviar sua resposta aos neurônios da camada  $F^{(1)}$ , i.e.

$$W_{ij}^{(b)} = 1, \text{ com } \begin{cases} i = 1, \dots, n \\ j = 1, \dots, m \end{cases}$$

- O parâmetro de vigilância  $\{\rho\}$  tem a missão de arbitrar o teste de similaridade a ser executado pelos neurônios da camada de entrada:
  - Seu valor deve estar compreendido entre 0 e 1 ( $0 < \rho < 1$ ).
  - Valores muito elevados** → Alta capacidade de discriminação de classes.
  - Valores muito baixos** → Pequena capacidade de discriminação, pois somente diferenças estruturais entre as amostras serão consideradas.

11

### 3. Aprendizizado da ART-1

#### Etapa (II) – Fase de Reconhecimento

- Consiste de tentar categorizar (enquadrar) um vetor de entrada apresentado dentro de uma daquelas classes já disponíveis em  $F^{(2)}$ .
- aquele neurônio da camada  $F^{(2)}$  que tiver a maior ativação (casamento), frente ao vetor de entrada  $\{x\}$ , será o vencedor.
- O grau de ativação do  $j$ -ésimo neurônio de  $F^{(2)}$  é produzido pela ponderação dos sinais da entrada pelos respectivos pesos, ou seja:

$$u_j = \sum_{i=1}^n W_{ji}^{(f)} \cdot x_i, \text{ com } j = 1, \dots, m$$

- Conseqüentemente, o neurônio vencedor  $\{k\}$  é então aquele que produzir o maior valor de ativação, representando aqui a classe candidata em que vetor de entrada  $\{x\}$  poderá ser inserido.
- O vencedor produz valor 1 em sua saída p/ indicar a classe candidata do vetor de entrada apresentado, ao passo que todos os outros neurônios de  $F^{(2)}$  produzirão valores iguais a 0, ou seja:

$$y_j = \begin{cases} 1, & \text{se } j = k \\ 0, & \text{se } j \neq k \end{cases}, \text{ com } j = 1, \dots, m$$

- Como o número de neurônios em  $F^{(2)}$  é dinâmica (neurônios são adicionados conforme necessidade) de se criar novas classes, infere-se que somente 1 neurônio estará presente quando da apresentação do primeiro vetor de entrada

12

### 3. Aprendizizado da ART-1

#### Etapa (III) – Fase de Comparação (I)

- O propósito aqui é verificar se a classe candidata representada pelo neurônio vencedor é a mais apropriada para receber o referido vetor.
- Para tal propósito, os resultados produzidos pelos neurônios de  $F^{(2)}$  são realimentados para  $F^{(1)}$  a fim de se fazer o teste de similaridade.
- Aquele neurônio da camada  $F^{(2)}$  que tiver a maior ativação (casamento), frente ao vetor de entrada  $\{x\}$ , será o vencedor.
- Como somente o neurônio vencedor  $k$  produziu valor 1 em sua saída, os valores que serão então realimentados de  $F^{(2)}$  para  $F^{(1)}$  são os próprios pesos deste vencedor  $k$  aos respectivos neurônios de  $F^{(1)}$ .
- Conforme definido antes, tanto os elementos da matriz  $W^{(b)}$  como do vetor  $x$  são binários, sendo que isto facilita o processo comparativo.
- A regra normalmente utilizada consiste de comparar o valor do parâmetro de vigilância com a razão  $R$ , que é obtida da quantidade de elementos unitários que sejam comuns (considerando  $W_{jk}^{(b)}$  e  $x_j$ ) em relação à quantidade de elementos unitários presentes no vetor  $x$ .

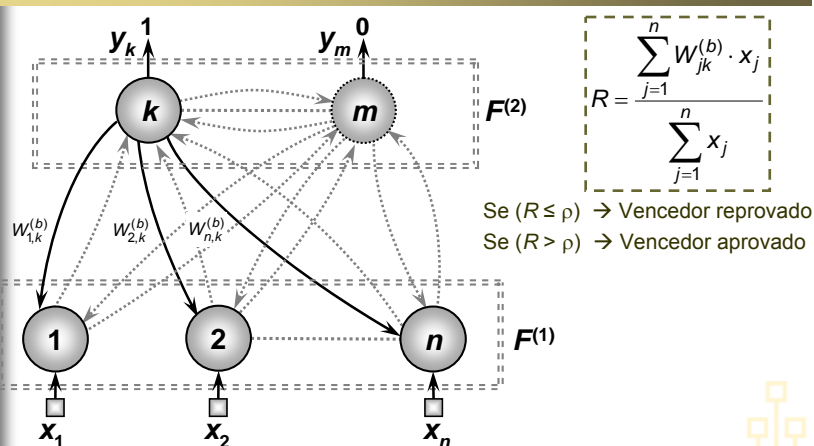
$$R = \frac{\sum_{j=1}^n W_{jk}^{(b)} \cdot x_j}{\sum_{j=1}^n x_j}$$

$\longrightarrow$  Quantidade de elementos unitários que sejam comuns.  
 $\longrightarrow$  Quantidade de elementos unitários presentes no vetor  $x$ .

13

### 3. Aprendizizado da ART-1

#### Etapa (III) – Fase de Comparação (II)



Exemplo:

$$W_{j,1}^{(b)} = \begin{bmatrix} j=1 & j=2 & j=3 & j=4 & j=5 & j=6 & j=7 & j=8 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}^T$$

$$x_j = \begin{bmatrix} j=1 & j=2 & j=3 & j=4 & j=5 & j=6 & j=7 & j=8 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^T$$

$R = 0.8$

14

### 3. Aprendizizado da ART-1

#### Etapa (IV) – Fase de Busca (I)

- Caso haja rejeição do neurônio vencedor frente ao teste de similaridade, buscar-se-á então aqui um outro neurônio (pertencente a  $F^{(2)}$ ) que tenha talvez mais afinidade com a amostra inserida.
- Para tanto, o vencedor do estágio anterior deverá ser desabilitado, a fim de se evitar que o mesmo possa novamente ganhar a competição.
- Uma das formas de se proceder tal operação está em atribuir o valor 0 ao potencial de ativação  $\{u_k\}$  daquele neurônio que era vencedor, pois assim um outro neurônio deverá agora o ganhador da competição.
- Desta forma, conforme se pode verificar no fluxograma do slide 10, procedem-se sucessivas vezes com a aplicação seqüencial das Fases (II), (III) e (IV) até que seja possível a obtenção de um neurônio vencedor de  $F^{(2)}$  que seja aprovado no teste de similaridade.
- Caso contrário, se todos os neurônios atualmente presentes em  $F^{(2)}$  forem reprovados no teste de similaridade, assumir-se-á então que o referido vetor de entrada esteja representando uma nova classe.
- Conseqüentemente, um novo neurônio será agora adicionado à  $F^{(2)}$ , a fim de representar aqui os atributos daquele referido padrão.
- De fato, tal neurônio irá perfeitamente incorporar as características deste padrão em seu vetor de pesos, pois esta será a 1ª amostra.

15

### 3. Aprendizizado da ART-1

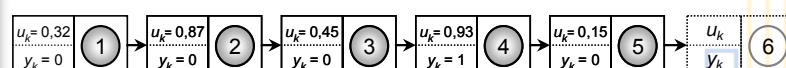
#### Etapa (IV) – Fase de Busca (II)

- Os procedimentos algorítmicos para esta adição de novos neurônios na camada  $F^{(2)}$  podem ser implementados de duas formas possíveis.

##### Primeira Forma: Estruturas de Dados Estáticas

Neurônio	1	2	3	4	5	6	7	...	50
Status	Ativo	Ativo	Ativo	Ativo	Ativo	Inativo	Inativo		Inativo
$u_k$	0,32	0,87	0,45	0,93	0,15	----	----		----
$y_k$	0	0	0	1	0	----	----		----

##### Segunda Forma: Estruturas de Dados Dinâmicas



16



### 3. Aprendizizado da ART-1

#### Etapa (V) – Fase de Atualização (I)

- Após a conclusão bem sucedida da Fase de Comparação (III), procede-se então com a fase de atualização dos pesos de  $\mathbf{W}^{(b)}$  e  $\mathbf{W}^{(f)}$ , a fim de incorporar o conhecimento trazido pela última amostra.

- Para tanto, somente aqueles elementos referentes ao neurônio vencedor  $\{k\}$  serão atualizados por meio das seguintes expressões:

$$W_{jk}^{(b)}(t+1) = W_{jk}^{(b)}(t) \cdot x_j, \text{ com } j = 1, \dots, n$$

$$W_{kj}^{(f)}(t+1) = \frac{W_{jk}^{(b)}(t) \cdot x_j}{\frac{1}{2} + \sum_{i=1}^n W_{ik}^{(b)}(t) \cdot x_i}, \text{ com } j = 1, \dots, n$$

- O ajuste dos pesos do neurônio vencedor em relação à matriz  $\mathbf{W}^{(b)}$  se processa por intermédio da realização de uma simples operação lógica de conjunção booleana (porta AND) entre os elementos  $W_{jk}^{(b)}$  e  $x_j$ .
- Já o ajuste dos pesos do neurônio vencedor em relação à matriz  $\mathbf{W}^{(f)}$  se processa por meio de duas operações.
  - Numerador** → Aplicada porta AND entre os elementos  $W_{jk}^{(b)}$  e  $x_j$ .
  - Denominador** → Quantidade de elementos unitários que estejam justapostos, tanto nos pesos do neurônio vencedor (em referência a  $\mathbf{W}^{(b)}$ ) como nas componentes do vetor de entrada.

17

### 3. Aprendizizado da ART-1

#### Etapa (V) – Fase de Atualização (II)

**EXEMPLO DE AJUSTE  $\mathbf{W}^{(b)}$**   $\left\{ W_{jk}^{(b)}(t+1) = W_{jk}^{(b)}(t) \cdot x_j, \text{ com } j = 1, \dots, n \right.$

		j=1	j=2	j=3	j=4	j=5	j=6	j=7	j=8
Pesos do vencedor	$W_{j,2}^{(b)}(t)$	1	0	0	1	1	1	0	1
Entradas da ART-1	$x_j$	1	0	1	1	1	0	1	1
Ajuste de $\mathbf{W}^{(b)}$ (Porta AND)	$W_{j,2}^{(b)}(t+1)$	1	0	0	1	1	0	0	1

**EXEMPLO DE AJUSTE  $\mathbf{W}^{(f)}$**   $\left\{ W_{kj}^{(f)}(t+1) = \frac{W_{jk}^{(b)}(t) \cdot x_j}{\frac{1}{2} + \sum_{i=1}^n W_{ik}^{(b)}(t) \cdot x_i}, \text{ com } j = 1, \dots, n \right.$

		j=1	j=2	j=3	j=4	j=5	j=6	j=7	j=8
Pesos do vencedor	$W_{j,2}^{(b)}(t)$	1	0	0	1	1	1	0	1
Entradas da ART-1	$x_j$	1	0	1	1	1	0	1	1
Ajuste de $\mathbf{W}^{(f)}$	$W_{2,j}^{(f)}(t+1)$	$\frac{2}{9}$	0	0	$\frac{2}{9}$	$\frac{2}{9}$	0	0	$\frac{2}{9}$

18

### 3. Aprendizizado da ART-1

#### *Etapa (VI) – Fase de Disponibilização*

- De acordo com o fluxograma do slide 10, a sexta etapa do processo de treinamento da rede ART-1 consiste de disponibilizar o rótulo da classe em que a amostra apresentada vai ser inserida.
- Neste caso, após a execução de todas as fases anteriores, basta-se aplicar as expressões seguintes a fim de obter a classe selecionada, pois será aquela representada pelo neurônio vencedor da competição de proximidade.

$$u_j = \sum_{i=1}^n W_{ji}^{(f)} \cdot x_i, \text{ com } j = 1, \dots, m$$

$$k = \arg \max_j \{u_j\}$$

$$y_j = \begin{cases} 1, & \text{se } j = k \\ 0, & \text{se } j \neq k \end{cases}, \text{ com } j = 1, \dots, m$$



19

### 4. Algoritmo da ART-1

#### *Fase de treinamento e operação*

**Início {Algoritmo ART-1 // Fase de Treinamento e Operação}**

```

<1> Apresentar o vetor de entradas { x };
<2> Inicializar as matrizes  $W^{(f)}$  e  $W^{(b)}$ ; {conforme (10.1) e (10.2)}
<3> Especificar o parâmetro de vigilância { ρ }; {definido entre 0 e 1}
<4> Repetir as instruções:
    {
        <4.1> Calcular as ativações neurais {  $u_j$  }; {conforme (10.3)}
        <4.2> Obter o neurônio vencedor atual {  $k$  }; {conforme (10.4)}
        <4.3> Calcular a razão de similaridade {  $R$  }; {conforme (10.6)}
        <4.4> Se (  $R > \rho$  ) fazer:
            {
                <4.4.1> Atualizar  $W^{(f)}$  e  $W^{(b)}$ ; {conforme (10.9) e (10.10)}
                <4.4.1> Atribuir  $x$  à classe do neurônio vencedor;
                <4.4.3> Vencedor ← "Aprovado";
            }
        <4.5> Se { (  $R \leq \rho$  ) e ( Existe neurônio disponível ) } fazer:
            {
                <4.5.1> Desabilitar neurônio vencedor atual {  $u_k = 0$  };
                <4.5.2> Vencedor ← "Reprovado";
            }
        <4.6> Se { (  $R \leq \rho$  ) e ( Não existe neurônio disponível ) } fazer:
            {
                <4.6.1> Incluir ou habilitar um novo neurônio;
                <4.6.2> Atualizar  $W^{(f)}$  e  $W^{(b)}$ ; {conforme (10.9) e (10.10)}
                <4.6.3> Atribuir  $x$  à classe do novo neurônio;
                <4.6.4> Vencedor ← "Aprovado";
            }
        Até que: Vencedor = "Aprovado";
    }
<5> Habilitar novamente todos os neurônios em  $F^{(2)}$ .

```

**Fim {Algoritmo ART-1 // Fase de Treinamento e Operação}**



20

## *Fim da Apresentação*

**EPC-12**  
**Data de Entrega: 25/06/2012**

