



Exacta

ISSN: 1678-5428

exacta@uninove.br

Universidade Nove de Julho

Brasil

Santos, Danilo; Bonaldo, Marco Aurélio; Araújo Alves de, Sidnei; Bueno, Valdir

Redes neurais artificiais aplicadas em análise de perfis profissionais

Exacta, núm. 2, novembro, 2004, pp. 91-104

Universidade Nove de Julho

São Paulo, Brasil

Disponível em: <http://www.redalyc.org/articulo.oa?id=81000207>

- ▶ Como citar este artigo
- ▶ Número completo
- ▶ Mais artigos
- ▶ Home da revista no Redalyc

redalyc.org

Sistema de Informação Científica

Rede de Revistas Científicas da América Latina, Caribe, Espanha e Portugal  
Projeto acadêmico sem fins lucrativos desenvolvido no âmbito da iniciativa Acesso Aberto

# Redes neurais artificiais aplicadas em análise de perfis profissionais\*

Danilo Santos

Bacharel em Ciência da Computação – UNINOVE.  
[dani.lo.santos@promon.com.br](mailto:dani.lo.santos@promon.com.br)

Marco Aurélio Bonaldo

Bacharel em Ciência da Computação – UNINOVE.  
[marcocabonaldo@uol.com.br](mailto:marcocabonaldo@uol.com.br)

Sidnei Alves de Araújo

Mestre em Engenharia Elétrica (Ênfase em Engenharia de Computação) – Mackenzie;  
Professor no curso de Ciência da Computação – UNINOVE.  
[saraugo@uninove.br](mailto:saraugo@uninove.br)

Valdir Bueno

Bacharel em Ciência da Computação – UNINOVE.  
[operacao@barsaplaneta.com.br](mailto:operacao@barsaplaneta.com.br)

## Resumo

Este trabalho explora a utilização de redes neurais artificiais auto-organizáveis (modelo de Kohonen) para a classificação de perfis profissionais, agrupando-os por áreas de atuação, com a finalidade de auxiliar no processo de triagem de candidatos, na área de Recursos Humanos (RH). O resultado prático dessa implementação é viável para a área de RH, tendo em vista que as redes neurais têm a capacidade de lidar com dados imprecisos, incompletos ou totalmente novos, adaptando-se rapidamente a eles.

## Palavras-chave

*Kohonen. Perfil Profissional. Redes Neurais.*

---

\* Trabalho de Conclusão de Curso (TCC), sob orientação de Sidnei Alves de Araújo. Ciência da Computação (2003) – UNINOVE.

SANTOS, Danilo; BONALDO, Marco Aurélio; ARAÚJO, Sidnei Alves de; BUENO, Valdir. Redes neurais artificiais aplicadas em análise de perfis profissionais. *Exacta*. v. 2, p. 91-104. São Paulo: UNINOVE, nov. 2004.



## Artificial neural networks applied to analysis of professional profiles

### Abstract

This work is aimed at exploring the utilization of artificial self organizing neural networks, Kohonen's model, for classification of professional profiles, with the finality of assisting the process of triaging candidates in Human Resources (HR) areas. The practical result of this implementation may be viable to areas in HR, knowing that the neural networks have the ability to get along and adapt rapidly to imprecise data, incomplete or totally new information.

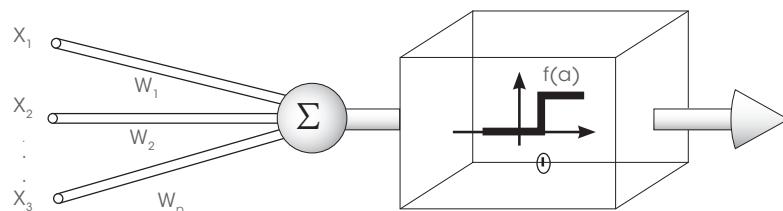
### Key words

*Kohonen. Neural networks. Professional profile.*

## Introdução

As redes neurais artificiais (RNAs) consistem em uma técnica de Inteligência Artificial (IA) que visa solucionar problemas baseando-se na estrutura do cérebro humano. Mediante um processo de treinamento, elas são capazes de aprender por meio da experiência (CARRÃO et al., 2002). Podemos dizer que suas principais características são: a forma de processamento distribuído, a habilidade de aprender e a capacidade de lidar com dados imprecisos (RUSSEL; NORVIG, 2004; KASABOV, 1998).

As pesquisas com redes neurais artificiais se iniciaram na década de 40, com o neurofisiologista Warren McCulloch e o matemático Walter Pitts, da Universidade de Illinois, os quais fizeram uma comparação entre células nervosas vivas e o processo eletrônico (FU, 1994; KOVÁCS, 2002). Essa comparação foi feita usando-se um modelo composto de resistores variáveis e amplificadores representando conexões sinápticas em um neurônio biológico (FERNANDES, 2003; RICH, 1993; RUSSEL; NORVIG, 2004). A Figura 1 ilustra o neurônio artificial projetado por McCulloch e Pitts.



**Figura 1 – Modelo de neurônio artificial**

Fonte: FERNANDES, 2003, p. 60.

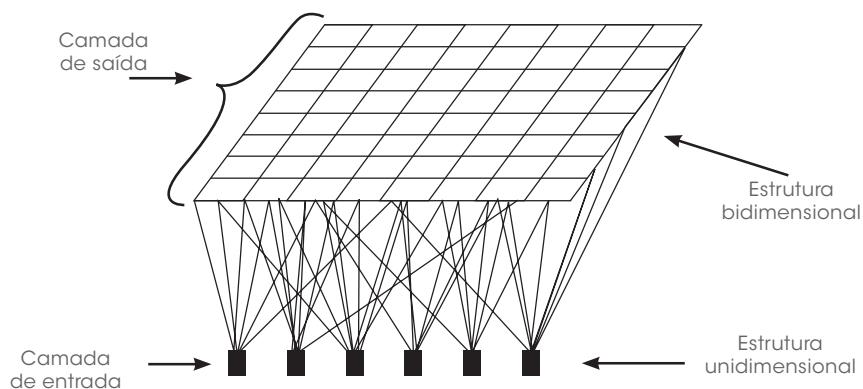
O primeiro modelo de rede neural, denominado *perceptron*, foi implementado por Frank Rosenblatt, em 1958. Trata-se de um modelo bastante limitado que pode ser utilizado apenas para classificação de padrões linearmente separáveis (HAYKIN, 2001; KOVÁCS, 1997; RICH, 1993; RUSSEL; NORVIG, 2004). Embora tenha suas limitações, o *perceptron* serviu de base para a criação de outros algoritmos de treinamento, tais como *Backpropagation* e de Kohonen, muito utilizados atualmente.

Atualmente, as redes neurais artificiais são utilizadas nas mais diversas áreas do conhecimento, tais como Engenharia, Economia, Agronomia e Recursos Humanos (RH), geralmente em problemas de reconhecimento e classificação de padrões. Há diversos modelos de RNAs, e cada um desses modelos é mais adequado para resolução de uma determinada classe de problemas. Algumas características que variam entre os modelos são: algoritmo de aprendizado, arquitetura da rede e a forma como são feitas as conexões entre as camadas.

O modelo explorado neste trabalho é um tipo de rede neural de aprendizado não supervisionado, conhecido como modelo de Kohonen, ou mapa auto-organizável – detalhado no item 1 –, que será implementado nesta pesquisa para classificar diferentes perfis profissionais, visando auxiliar a área de RH de empresas, na atividade de triagem de currículos.

## 1. Modelo de Kohonen

O modelo de Kohonen, cujo aprendizado é do tipo não supervisionado, possui apenas duas camadas (entrada e saída). Esse tipo de rede pode apresentar diferentes formatos geométricos em sua camada de saída (CABRAL JUNIOR; TIMOSZCZUK, 2003; HAYKIN, 2001); no entanto, o formato mais utilizado é o retangular. A Figura 2 ilustra um exemplo de arquitetura do modelo.



**Figura 2 - Exemplo de arquitetura do modelo de Kohonen.**

Fonte: Elaboração própria.

A rede de Kohonen é caracterizada pela formação de um mapa topográfico dos padrões de entrada, no qual as localizações espaciais dos neurônios da grade de saída são indicativas das características estatísticas intrínsecas contidas nos padrões de entrada, o que justifica o nome ‘mapa auto-organizável’ (KOHONEN, 1990). Esse modelo de rede neural foi inspirado no córtex sensitivo, em que regiões vizinhas respondem por partes sensitivas localizadas próximas do corpo humano (RITTER; KOHONEN, 1989). Como suas principais características, podemos citar:

- 1) Competição – para cada padrão de entrada apresentado à rede, os neurônios da camada de saída competem entre si para decidir quem será o vencedor. Esse neurônio vencedor representará o conjunto de características descrito pelo padrão de entrada em questão;
  - 2) Cooperação – o neurônio vencedor determina a localização espacial de uma vizinhança topológica de neurônios que devem ser excitados, fornecendo, assim, a base para a cooperação entre os vizinhos;
  - 3) Generalização – capacidade de a rede neural fornecer respostas adequadas a um padrão de entrada que não foi aprendido previamente.

## 1.1. Treinamento do modelo de Kohonen

Na fase de treinamento, os neurônios da camada de saída competem entre si para que haja um vencedor a cada nova entrada do conjunto de treinamento, ou seja, sempre que é apresentado um padrão do conjunto de treinamento à rede neural, existe uma competição entre os neurônios da camada de saída para representar a entrada naquele momento (BRAGA, 2000). Uma particularidade do modelo de Kohonen é que os pesos dos neurônios vizinhos do vencedor também são atualizados, configurando o processo de cooperação. No início do treinamento, as vizinhanças de atualização são grandes e podem abranger toda a grade de saída. Durante o treinamento, a redução do tamanho da vizinhança favorece uma seleitividade entre os padrões de entrada e a formação de regiões de maior atividade em torno do neurônio vencedor. Esse é o processo responsável pela formação de mapas que conservam a topologia (KOHONEN, 1990). O sucesso no processo de aprendizagem da rede depende não somente do número de iterações, mas também de uma melhor representação dos padrões de entrada e dos parâmetros do algoritmo, tais como a taxa de aprendizagem e a função de vizinhança. É importante ressaltar que esses parâmetros são obtidos de forma empírica.

Para um melhor entendimento, mostramos, a seguir, os seis passos necessários à implementação do algoritmo de treinamento do modelo de Kohonen (CABRAL; TIMOSZCZUK, 2003; HAYKIN, 2001):

- 1) Antes de qualquer operação na rede, inicializam-se todos os pesos ( $\vec{w}_{ij}$ ) das sinapses com valores aleatórios pequenos em relação aos valores de entrada ( $\vec{x}_j$ );
- 2) Apresenta-se uma nova entrada ( $\vec{x}_j$ ) (do conjunto de treinamento);
- 3) Calculam-se as distâncias entre os neurônios da camada competitiva e os da camada de entrada. Para cada neurônio da camada competitiva, calcular a distância  $d_i$  entre o neurônio  $\vec{w}_{ij}$  da camada competitiva e cada neurônio  $\vec{x}_j$  da camada de entrada.

$$d_i(t) = \sqrt{\sum_{j=1}^J (x_j(t) - w_{ij}(t))^2}$$

onde:

$j = 1, \dots, J$  – Neurônios de entrada  
 $i = 1, \dots, N$  – Neurônios de saída

$x_j(t)$  = valor do neurônio de entrada  $j$  no tempo  $t$ .

$w_{ij}(t)$  = peso sináptico no tempo  $t$ .

- 4) Seleciona-se o neurônio vencedor: seleção do neurônio  $\vec{w}_{ij}$ , cuja distância euclidiana  $d_i$  seja a menor possível, ou seleção do neurônio  $i$ , cujo vetor de pesos sinápticos  $\vec{w}_{ij}$  é o mais parecido com o vetor de entrada  $\vec{x}_j$ , sendo:

$$d_{i^*} < d_i \quad i = 1, \dots, N; i \neq i^*$$

- 5) Atualizam-se, no próprio neurônio vencedor, seus vetores de pesos e os de sua vizinhança (numa área que abrange  $\Delta(t)$  neurônios à sua volta), obedecendo ao seguinte cálculo:

$$w_{ij}(t+1) = w_{ij}(t) + n(t) \cdot (x_j(t) - w_{ij}(t))$$

no qual:

$n(t)$  = taxa de aprendizagem que deve ser diminuída no decorrer do treinamento.

$\Delta(t)$  = número de neurônios vizinhos afetados pela atualização de pesos.

6) Volta-se ao passo 2 com um novo conjunto de treinamento.

## 2. Perfil profissional

O perfil profissional é um conjunto de características que precisam ser encontradas em um candidato para que este possa ocupar um determinado cargo. Esse perfil pode ser dividido em duas partes: conhecimentos técnicos e perfil comportamental, em que se observam as diversas habilidades do candidato, conforme ilustrado na Tabela 1. Neste trabalho, consideraremos apenas o perfil técnico, em razão da subjetividade e dificuldade encontradas na mensuração dos valores que expressam o perfil comportamental.

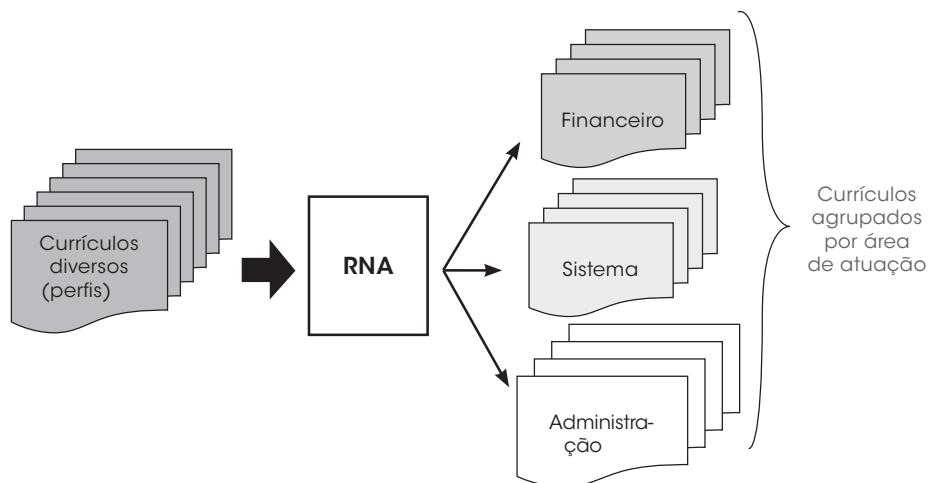
**Tabela 1 – Perfil profissional.**

<b>Conhecimentos técnicos</b>	<b>Perfil comportamental</b>
<ul style="list-style-type: none"><li>• escolaridade</li><li>• área de formação</li><li>• idiomas</li><li>• experiência profissional</li><li>• informática</li></ul>	<ul style="list-style-type: none"><li>• comunicação</li><li>• relacionamento interpessoal</li><li>• bom senso</li><li>• postura</li><li>• ética</li></ul>

Fonte: Elaboração própria.

Devido à grande competição e à necessidade de encontrar profissionais cada vez mais qualificados e que atendam aos perfis solicitados, as empresas acabam recorrendo a *sites* de busca de profissionais, ou criando um local em seu próprio *site*, em que os profissionais possam cadastrar seus currículos. Quando a empresa cadastrá uma vaga num desses *sites* de busca, o número de currículos recebidos é muito grande, perdendo-se muito tempo em sua triagem, já que, apesar do avanço alcançado por essas empresas neste segmento, esses sistemas ainda são muito falhos, pois possibilitam ao candidato enviar seu currículo para qualquer vaga

cadastrada, ou o restringem a apenas uma delas. Por esse motivo, resolvemos implementar um algoritmo de RNA na tentativa de minimizar esse tipo de problema. Essa implementação tem como objetivo auxiliar a área de RH na tarefa de classificação e triagem de currículos. A idéia de como deve ocorrer essa classificação é ilustrada na Figura 3.



**Figura 3 – Classificação de currículos por área de atuação.**

Fonte: Elaboração própria.

### 3. Modelagem

Para que as informações referentes aos perfis profissionais pudessem servir de entrada para treinamento da RNA, realizamos os seguintes procedimentos:

Escolhemos 15 perfis profissionais entre as áreas administrativas de saúde, tecnologia, direito e letras. Esses perfis foram analisados e tiveram seus atributos quantificados para compor o conjunto de treinamento da rede. Após diversos experimentos e com base nos estudos de Araújo (2002), descartamos os atributos de menor desvio padrão e selecionamos os sete principais de maior desvio padrão, fundamentais para definir a função de um profissional. Esses atributos foram valorados com pesos diferentes, de acordo com seu grau de importância. Fizemos ainda uma normalização dos valores de entrada, deixando-os na faixa entre 0 e 1, por se tratar de uma rotina de praxe em RNAs.

Na Tabela 2, é possível visualizar os perfis que compõem o conjunto de treinamento com seus respectivos atributos já quantificados e normalizados.

**Tabela 2 - Conjunto de treinamento.**

Perfis (cargos)	Atributos						
	Escolaridade	Área de formação	Experiência na função	Área de atuação	Atividades desenvolvidas	Inglês	Espanhol
1. Análise de sistemas	1	0,35	0,4	0,35	0,35	0,9	0,4
2. Análise de suporte	0,7	0,36	0,2	0,35	0,35	0,5	0,4
3. Engenharia civil	1	0,31	0,2	0,31	0,31	0,9	0,4
4. Nutricionista	1	0,69	0,4	0,69	0,69	0,3	0
5. Médico	1	0,66	0,6	0,66	0,66	0,9	0,6
6. Secretária	0,7	0,18	0,4	0,12	0,19	0,3	0,4
7. Secretária bilíngüe	0,7	0,18	0,4	0,12	0,19	0,9	1
8. Tradutor de inglês	1	0,17	0,4	0,19	0,19	0,9	0,4
9. Tradutor de espanhol	1	0,17	0,4	0,19	0,19	0,3	1
10. Advogado tributário	1	0,11	0,6	0,11	0,11	0,3	0,4
11. Revisor de trad. jurídica	1	0,11	0,4	0,11	0,11	0,9	1
12. Auxiliar administrativo	0,5	0	0,2	0,12	0,12	0,3	0,4
13. Analista financeiro	1	0,15	0,4	0,16	0,16	0,3	0,4
14. Analista de RH	1	0,12	0,4	0,14	0,14	0,3	0,4
15. Controller	1	0,19	0,6	0,15	0,15	0,7	0,4

Fonte: Elaboração própria

O conjunto de profissões descritas na Tabela 2 é utilizado para treinamento da rede. O propósito é que, depois desse treinamento, tomando-se o perfil de um candidato (currículo) e apresentando-o à rede, esta seja capaz de classificá-lo corretamente em uma das áreas formadas no mapa de saída, de acordo com suas características intrínsecas. Dessa forma, é possível saber de quais cargos esse perfil mais se aproxima.

## 4. Implementação e parametrização da rede

A rede neural para classificação dos perfis profissionais foi implementada utilizando a linguagem de programação Delphi 5.0. Desenvolveu-se uma aplicação genérica para classificação de padrões, aqui utilizada para reconhecimento e classificação de perfis profissionais. Essa aplicação é composta, basicamente, de três módulos: 1. Módulo de parametrização da rede neural, 2. Módulo de aquisição dos padrões de treinamento e 3. Módulo de treinamento e visualização do mapa de saída.

Nesta aplicação, utilizamos os seguintes parâmetros:

- 1) Arquitetura da rede: sete neurônios na camada de entrada e 169 neurônios na camada de saída, dispostos numa grade retangular (13 x 13);
- 2) Taxa de aprendizagem:  $\eta(t) = 0,2$ ;
- 3) Equação de aprendizagem: define a taxa real de aprendizagem do vetor de pesos do neurônio, de acordo com sua posição na grade de saída. Quanto mais longe do vencedor, menor é a taxa. Usamos as seguintes equações:

$$EQ1 = \frac{n(t)}{(1 + (\frac{\Delta}{5}))}$$

$$EQ2 = n(t) \cdot (0,9 + (\frac{1}{2^\Delta}) \cdot (\frac{1}{10})), \text{ onde:}$$

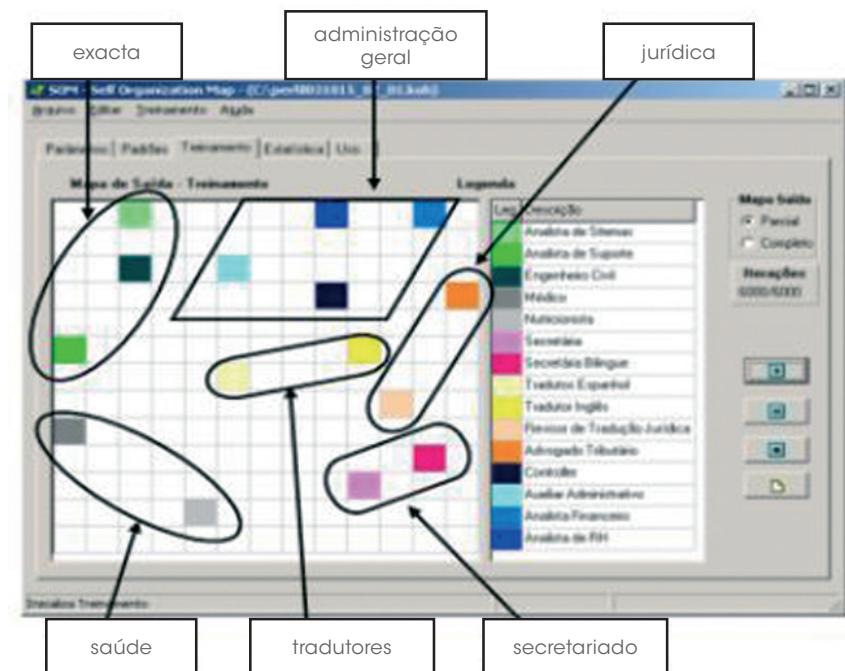
$n(t)$  = Valor da taxa de aprendizagem utilizada no treinamento.  
 $\Delta$  = Raio da vizinhança a ser atualizado.

- Quantidade de iterações: 6 mil.

## 5. Resultados obtidos

Depois de incluído o conjunto de padrões e parametrizada a rede, inicia-se o processo de treinamento. Nesta fase, os neurônios da camada de saída devem competir entre si, representando o neurônio vencedor de cada iteração um perfil profissional do conjunto de treinamento. Perfis

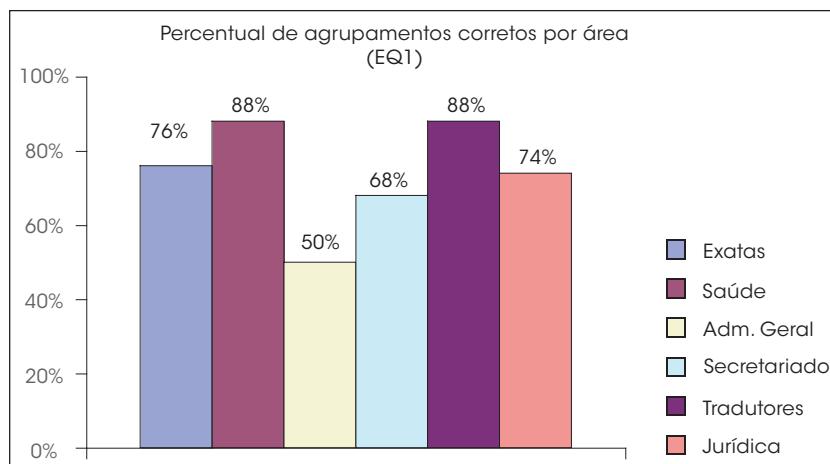
com características semelhantes tendem a ser representados por neurônios vizinhos na camada de saída (BARRETO, 2001). Isso leva à formação do mapa topográfico de características. Um dos mapas obtidos durante os treinamentos pode ser visto na Figura 4.



**Figura 5 – Exemplo de um mapa de saída após treinamento da rede**  
Fonte: Elaboração própria.

Para averiguar a eficiência da rede na formação dos agrupamentos dos diferentes perfis profissionais, fizemos duas sessões de treinamentos, utilizando diferente equação de aprendizagem em cada uma delas.

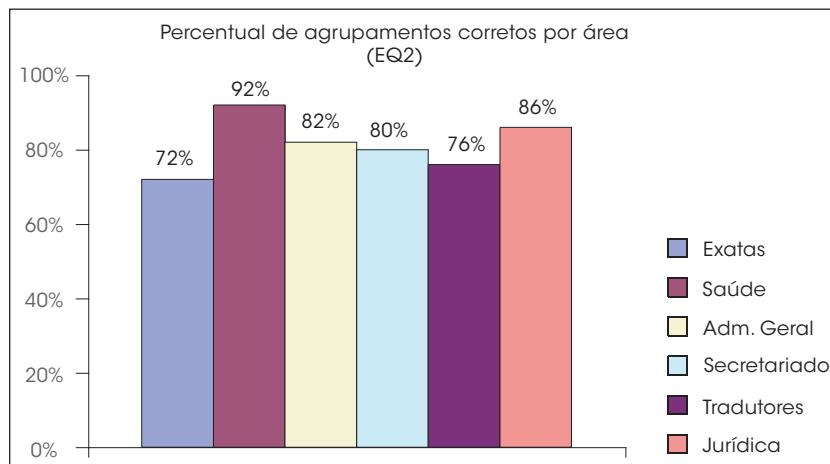
Inicialmente, realizamos 50 treinamentos utilizando os parâmetros descritos no ‘4. implementação e parametrização da rede’, considerando a 1<sup>a</sup> equação de aprendizagem (EQ1). O gráfico ilustrado na Figura 5 sintetiza o resultado dessa sessão de treinamentos.



**Figura 5 - Gráfico dos percentuais de agrupamentos por área de atuação em 50 treinamentos, considerando a equação de aprendizagem EQ1.**

Observa-se, na da Figura 5, que a maior parte dos agrupamentos ficou abaixo de 80%. Assim, consideramos insatisfatório esse resultado e concluímos que a equação para atualização de pesos EQ1, utilizada com os demais parâmetros da rede, mostrou-se ineficiente.

Procedemos, então, a mais 50 treinamentos, utilizando a 2<sup>a</sup> equação de aprendizagem (EQ2), mantendo-se inalterados os demais parâmetros da rede, e analisamos a quantidade de agrupamentos corretos, na camada de saída, em cada um dos treinamentos. A Figura 6 ilustra, de forma gráfica, um resumo estatístico desses treinamentos.



**Figura 6 - Gráfico dos percentuais de agrupamentos por área de atuação em 50 treinamentos, considerando a equação de aprendizagem EQ2.**

Observando a Figura 6, nota-se que, na maioria das áreas, o percentual de agrupamentos corretos foi superior, se comparado com o percentual mostrado no gráfico da Figura 5. Além disso, a média geral de agrupamentos (81%) também foi superior à demonstrada no gráfico anterior (74%). Entretanto, duas das seis áreas ficaram com o percentual de agrupamento ainda insatisfatório (abaixo de 80%): trata-se das áreas de Tradutores e Exatas em que os agrupamentos corretos ocorreram, respectivamente, em 76% e 72% dos casos. Isso pode ter sido causado pela forma como foram quantificadas as características que representam os perfis dessas áreas.

## Conclusão

Neste trabalho, implementamos uma rede neural para classificar e agrupar perfis profissionais semelhantes, com o intuito de contribuir para o processo de triagem de currículos na área de RH. Dessa implementação, conclui-se que a representação do conhecimento é uma etapa de vital importância para resolução de problemas usando redes neurais artificiais. Um problema mal representado pode inviabilizar um projeto de RNA.

O percentual médio de agrupamentos obtido na segunda sessão de treinamentos (81%), considerado satisfatório, pode ser aperfeiçoado desde que seja feita uma melhor análise dos atributos que compõem um perfil.

Outra observação que podemos fazer é que há perfis de diferentes áreas com atributos bastante parecidos. Isso é decorrente do processo de representação do conhecimento.

É válido lembrar que outros fatores, tais como taxa de aprendizagem, arquitetura da rede e quantidade de iterações, influem diretamente na formação do mapa de saída. Esses parâmetros são escolhidos de forma empírica e podem ser melhorados mediante análise mais criteriosa e um maior número de experimentos.

Contudo, pelos resultados aqui apresentados, concluímos que a rede foi capaz de fazer os agrupamentos propostos inicialmente, mesmo não tendo atingido percentuais tão expressivos.



Para que a aplicação possa ser de fato utilizada, torna-se necessário fazer os ajustes aqui mencionados e treinar a rede com um maior número de padrões. Além disso, deve-se implementar um módulo interface para que o usuário final possa utilizar essa aplicação.

## Referência

- ARAÚJO, Sidnei Alves de. *Classificação de perfis profissiográficos usando a rede de Kohonen*. 2002. 107 p. Dissertação (Mestrado em Engenharia Elétrica). São Paulo, Universidade Presbiteriana Mackenzie, 2002.
- BARRETO, Jorge Muniz. *Inteligência artificial: no limiar do século XXI*. 3. ed. Florianópolis: Duplic, 2001.
- BRAGA, Antônio de Pádua. *Redes neurais artificiais: teoria e aplicações*. Rio de Janeiro: LTC, 2000.
- CABRAL JUNIOR, Euvaldo Ferreira; TIMOSZCZUK, A. P. *Redes neurais artificiais: modelos em C*. São Paulo: Suprema, 2003.
- CARRÃO, A. C.; FARIA, P. C.; MENDONÇA, R. R.; ARAÚJO, S. H. *Redes neurais artificiais aplicadas em análise de crédito*. 2002. 46 p. Trabalho de Conclusão de Curso (TCC) em Ciência da Computação. Centro Universitário Nove de Julho (UNINOVE). São Paulo, 2002.
- FERNANDES, Anita Maria da Rocha. *Inteligência artificial: noções gerais*. Florianópolis: Visual Books, 2003.
- FU, LiMin. *Neural networks in computer intelligence*. New York: McGraw-Hill, 1994.
- HAYKIN, Simon. *Redes neurais: princípios e prática*. Trad. Paulo Martins Engel. 2. ed. Porto Alegre: Bookman, 2001.
- KASABOV, Nikola K. *Foundations of neural networks, fuzzy systems and knowledge engineering*. 2. ed. Massachusetts: The MIT Press, 1998.
- KOHONEN, T. The self-organizing map. *Proceedings of the Institute of Electrical and Electronics Engineers*. v. 78, p. 1464-1480, 1990.
- KOVÁCS, Zsolt Lászlo. *O cérebro e a sua mente: uma introdução à neurociência computacional*. São Paulo: Acadêmica, 1997.
- \_\_\_\_\_. *Redes neurais artificiais: fundamentos e aplicações*. 3. ed. São Paulo: Editora Livraria da Física, 2002.
- RICH, Elaine. *Inteligência artificial*. 2. ed. São Paulo: Makron Books, 1993.
- RITTER, Helge; KOHONEN, Teuvo. Self-organizing semantic maps. *Biological Cybernetics*. v. 61, p. 241-254, 1989.
- RUSSEL, Stuart; NORVIG, Peter. *Inteligência artificial*. Rio de Janeiro: Elsevier, 2004.