

Estudo de relações de similaridade difusas e sistemas neuro-fuzzy aplicados ao raciocínio baseado em casos

Jonas Henrique Mendonça¹, Sandra Sandri¹, Flávia Martins-Bedê¹, Ricardo Guimarães², Omar Carvalho³

¹Instituto Nacional de Pesquisas Espaciais
12201-970, São José dos Campos, SP
jonas.henrique01@gmail.com, sandri@lac.inpe.br, toledomartins@gmail.com

²Laboratório de Geoprocessamento / Instituto Evandro Chagas / SVM / MS
66093-971, Belém, PA
ricardojpsg@gmail.com

³Centro de Pesquisas René Rachou / FIOCRUZ
30190-002, Belo Horizonte, MG
omar@cpqrr.fiocruz.br

Abstract. This work is a study of an approach of case-based reasoning using fuzzy similarity relations weighted. We used several metrics to obtain the weight vectors that would later be used in data classification tasks. The main objective of this work is to show the pros and cons of relationships used. We use as a case study a problem from the real world to estimate the prevalence of schistosomiasis in a region of Brazil. The results for this test are based, in most cases, superior to approaches previously applied.

Keywords: Fuzzy logic, similarity relations, fuzzy-ART

1 Introdução

A técnica de reconhecimento baseada em casos (RBC) [6] se propõe a resolver um problema usando um princípio que pode ser declarado como “problemas semelhantes tem soluções semelhantes” [1]. A base consiste em problemas resolvidos modelados como pares (problema, solução) e é utilizada para determinar a solução para um novo problema. O primeiro passo deste procedimento consiste em recuperar problemas na base que são semelhantes ao problema considerado: ela determina os casos da base que são relevantes para a solução do problema. A segunda etapa consiste em reutilizar as soluções desses problemas relevantes, adaptando-os para o problema considerado. Pesos podem ser anexados aos casos, de maneira que os considerados mais importantes para uma determinada aplicação tenham pesos mais elevados. Vetores de pesos também podem ser associados às variáveis de descrição do problema. Pode-se

usar o vetor de pesos iguais para todos os casos o que recai em ponderar as variáveis independentemente dos casos. Mas, pode-se atribuir vetores de pesos individuais para cada caso, de modo que mais atributos significativos no interior de um caso recebam pesos mais elevados. Em [3] e [8], relações de similaridade difusa associadas a cada descrição e solução no espaço de variáveis foram utilizadas para derivar vetores de pesos individuais através do algoritmo de aprendizagem proposto em [14]. Foi demonstrado que o uso de vetores de pesos individuais para cada caso do conjunto de treinamento tende a levar a melhores resultados que a utilização de casos não ponderados.

O problema com a utilização dos pesos é que o processo de aprendizado geralmente é computacionalmente caro. O que pode impossibilitar sua utilização em grandes bases de casos. Uma abordagem para permitir a aprendizagem ponderada em bases de dados grandes consiste na extração de fragmentos da base de dados e na obtenção de pesos para cada um desses fragmentos. O cálculo da solução para um novo caso utiliza os fragmentos cujos problemas são semelhantes aos casos em questão. Um modelo para fragmentar as bases foi proposto em [5]. O método proposto é baseado na similaridade binária entre casos, chamado relação de semelhança entre casos (CRR) que leva em conta tanto a semelhança nas entradas quanto nas saídas do problema [12]. Estamos interessados no processo de aprendizagem ponderado para agrupamentos gerados por um dado CRR. Uma questão relevante é qual deve ser a base de treinamento para se obter os pesos dos casos em um agrupamento. Em [11] foi verificado que utilizar um conjunto maior de casos que somente àqueles do agrupamento como base de treinamento produz melhores resultados.

O objetivo deste trabalho é apresentar um estudo sobre as relações de similaridade que podem ser aplicadas a base de casos e como encontrar fragmentos da base de casos que mais se assemelham e possam gerar melhores resultados na tarefa de classificação.

Este trabalho está organizado da seguinte maneira: a seção 2 apresenta as definições das técnicas empregadas neste trabalho, a metodologia utilizada no processo de classificação é mostrada na seção 3, a seção 4 mostra um estudo de casos e os resultados obtidos. Por fim, as conclusões obtidas são apresentadas na seção 5.

2 Definições básicas e notações

Nesta seção apresentamos as definições e notações utilizadas neste trabalho.

2.1 Lógica difusa e relações de similaridade

Na lógica difusa [16], um operador $T : [0,1]^2 \rightarrow [0,1]$ é chamado uma *t-norma* se é comutativo, associativo, monotônico e seu elemento neutro é 1. Um operador $\perp : [0,1]^2 \rightarrow [0,1]$ é chamado uma *t-conorma* se é comutativo, associativo e monotônico com elemento neutro igual a 0. Podemos citar as operações; mínimo e produto como exemplo, de *t-normas* e máximo e soma limitada como exemplo de *t-conormas*.

Dada uma t-norma T , um operador de implicação residual é definido como $a \rightarrow b = \sup_{c \in [0,1]} T(a, c) \leq b$. Alguns exemplos muito conhecidos incluem:

- implicação Gödel, resíduo de $T = \min$, definido como $a \rightarrow b = 1$ se $a \leq b$ e $a \rightarrow b = b$ caso contrário;
- implicação de Goguen, definido como $a \rightarrow b = 1$ se $a \leq b$ e $a \rightarrow b = b/a$ caso contrário.

Uma relação de similaridade difusa S em um domínio A é um mapeamento $S : S : \Omega \times \Omega \rightarrow [0,1]$, que atribui a cada par ordenado (w, w') de elementos de A um valor que mede o quanto w e w' são similares. Estas relações foram originalmente introduzidas por Zadeh [17] como uma generalização da definição clássica de relações de equivalência.

Segundo [10] e [4], pode-se dizer que uma relação binária difusa S em um universo A é uma relação de similaridade em A se as seguintes propriedades de simetria, reflexividade e T-transitividade forem satisfeitas.

Alguns autores definem relações de similaridade como sendo simétricas e reflexivas, sem necessariamente obedecer a transitividade. Neste trabalho utilizamos esta definição.

2.2 Rede Fuzzy-ART

As redes Fuzzy-ART são modelos formalmente parecidos com as redes ART, porém realiza duas operações adicionais; uma relacionada à aprendizagem e outra ao pré-processamento das entradas. Uma rede Fuzzy-ART gera agrupamentos de vetores de características difusos (*fuzzy*). Mais especificamente, segundo [13], cada componente do vetor de entrada i é um valor de pertinência da função membro de uma determinada característica difusa, indicando o quanto esta característica está presente na amostra. Assim, a dinâmica de um sistema Fuzzy-ART é descrita em termos das operações da teoria de conjuntos nebulosos. O aprendizado sempre converge porque todos os pesos adaptativos são monotonicamente decrementados. Porém, sem uma etapa de pré-processamento das entradas, esta estabilidade poderia levar à proliferação das categorias quando os pesos decaíssem em zero. Este pré-processamento é chamado de codificação por complemento, isto é, as entradas passam a ser representadas por suas porções ligadas (força de presença dos atributos) e por suas porções desligadas (força de ausência dos atributos). A codificação por complemento normaliza os vetores de entrada e ao mesmo tempo preserva a amplitude das ativações individuais das características (cada componente do vetor de entrada). As computações dos operadores nebulosos E e OU são implementados por meio das funções de mínimo (\wedge) e máximo (\vee), respectivamente, conforme demonstrado na Tabela 1.

Tabela 1 - Analogia entre as redes ART e Fuzzy-ART

Fonte: [9]

	ART (BINÁRIO)	FUZZY-ART
Escolha de Categoria	$T_j = \frac{ I \cap W_j }{\alpha + W_j }$	$T_j = \frac{ I \wedge W_j }{\alpha + W_j }$
Critério de Similaridade	$\frac{ I \cap W_j }{ I } \geq \rho$	$\frac{ I \wedge W_j }{ I } \geq \rho$
Aprendizado Rápido	$W_j^{novo} = I \cap W_j^{velho}$	$W_j^{novo} = I \wedge W_j^{velho}$

onde I são os vetores de entrada, W_j são os vetores de pesos adaptativos e $\rho \in [0,1]$ é o parâmetro de vigilância.

2.3 Hipergrafos

Um hipergrafo é uma generalização de um grafo não-direcionado, onde arestas podem conectar qualquer número de vértices. Formalmente, isto pode ser representado como um par, $H=(N, E)$, onde N é um conjunto de vértices e E é um conjunto de subconjuntos não vazios de N chamado hiperarestas. O conjunto de hiperarestas E é então um subconjunto de $2^N \setminus \emptyset$, onde 2^N é o conjunto potência de N . Um “grafo comum” é então um hipergrafo no qual todas hiperarestas tem no máximo 2 elementos.

Pode-se salientar que cada grafo pode ser associado a um hipergrafo cujas hiperarestas são os cliques de um grafo inicial. Por outro lado, dado um hipergrafo $H=(N, E)$, uma hiperaresta $A \in E$ é dito máxima quando $\nexists B \neq A \in E, A \subset B$. Cada hiperaresta em E é um clique, portanto, o conjunto de hiperarestas máximo é o conjunto de cliques máximo de E .

2.4 Raciocínio baseado em casos

A técnica de Raciocínio Baseado em Casos (RBC), do inglês Case-based reasoning, consiste em utilizar um conjunto de soluções anteriores, com ou sem alterações dentro de um determinado domínio, para solucionar novos problemas [2]. De acordo com [1], o RBC, de uma forma generalizada pode ser dividido nas seguintes etapas:

- recuperação dos casos mais semelhantes;
- reutilização da informação e conhecimento do caso para resolver o problema;
- revisão da solução proposta, podendo ser adaptada a uma nova realidade;
- retenção de partes da experiência obtida para ser utilizada na resolução de novos problemas;

Segundo [15], o entendimento da técnica de RBC está implícito em assumir alguns princípios da natureza do mundo:

- Regularidade: o mundo é na maioria das vezes regular, as ações executadas nas mesmas condições tendem a ter os mesmos, ou similares resultados. Consequentemente, soluções para problemas similares são utilizáveis para o início da resolução de outros.
- Tipicidade: os tipos de problemas tendem a se repetir; as razões para as experiências são provavelmente as mesmas para as futuras ocorrências;
- Consistência: pequenas mudanças ocorridas no mundo requerem apenas mudanças na maneira como interpretamos o mundo, e consequentemente, pequenas mudanças nas soluções de novos problemas;
- Facilidade de adaptação: as coisas não se repetem exatamente da mesma maneira; as diferenças tendem a ser pequenas e pequenas diferenças são fáceis de compensar.

A qualidade de um sistema RBC depende do número de casos relevantes que farão parte da base de casos.

3 Metodologia

A metodologia aplicada neste trabalho foi originalmente proposta em [11]. Porém, como neste trabalho estamos tratando das diferentes relações de similaridades e métricas para calcular agrupamentos e resultados finais que podem ser aplicados, mostraremos as diferentes funções que foram utilizadas em cada uma das etapas.

3.1 Definições básicas:

Um caso c é definido como um par ordenado $c = (p, o) \in P \times O$ onde p é a descrição do problema e o a descrição de sua solução e $P = \{P_1 \times \dots \times P_n\}$ e O são respectivamente a descrição do problema e o espaço de soluções.

Seja $S_{out} \subseteq O^2$ uma relação de similaridade em O e seja $S_{in} \subseteq P^2$ uma relação de similaridade definida sobre todo o espaço do problema. S_{in} pode ser obtido usando uma função de agregação aplicada a um conjunto de relações de similaridade $\{S_1, \dots, S_n\}$ correspondente as variáveis que descrevem o problema. Por exemplo, usando média aritmética para S_{in} temos $S_{means}(p_i, p_j) = \frac{1}{n} \sum_{k=1}^n S_k(p_{ik}, p_{jk})$.

3.2 Casos de relações de semelhança difusa:

Considerando \emptyset um operador de implicação residual e definindo um *caso de relação de semelhança difusa* (FCRR) como uma relação $F_\emptyset : C^2 \rightarrow [0,1]$ definida como na equação 1.

$$F_\emptyset(c_a, c_b) = \begin{cases} 0, & \text{se } S_{in}(p_a, p_b) = 0 \\ \emptyset(S_{in}(p_a, p_b), S_{out}(o_a, o_b)), & \text{senão} \end{cases} \quad (1)$$

A relação F_\emptyset é uma formalização gradual do princípio básico do RBC: “problemas semelhantes tem soluções semelhantes”.

O resultado obtido nesta etapa do processo de classificação é usado para se calcular os agrupamentos correspondentes a base de casos. Além da relação apresentada na equação 1 pode-se usar relações com Goguen e seus respectivos cortes de nível como mostrado na equação 3.2.

$$\forall \alpha \in (0,1], F_{\emptyset,\alpha}(c_i, c_j) = \begin{cases} 1, & \text{se } F_\emptyset(c_i, c_j) \geq \alpha \\ 0, & \text{caso contrário} \end{cases} \quad (2)$$

3.3 Obtendo agrupamentos para uma base de casos:

Foram estudadas duas maneiras de se encontrar os agrupamentos para uma base de casos.

Na primeira delas, considera-se dois casos $c_a = (p_a, o_a)$ e $c_b = (p_b, o_b)$ em C . Seja R uma CCRR obtida como um corte de nível α para um dado F_\emptyset .

Baseado em uma CCRR R , um conjunto de casos pode ser organizado através de uma decomposição em agrupamentos baseada em uma relação de semelhança clássica. Muitos agrupamentos podem ser obtidos de R e podem ser representados como um hipergrafo. Mais precisamente, um hipergrafo $H = (C, E), E \subseteq C^2$ é dito compatível com a CCRR R se, e somente se, obedece as seguintes condições:

$$\begin{aligned} \forall c_a, c_b \in C, \text{ se } R(c_a, c_b) = 1, \text{ então } \exists h \in E \text{ tal que } \{c_a, c_b\} \subseteq h \\ \forall c_a, c_b \in C, \text{ se } R(c_a, c_b) = 0, \text{ então } \nexists h \in E \text{ tal que } \{c_a, c_b\} \subseteq h \end{aligned}$$

Na segunda, utilizou-se redes neurais artificiais (RNA) Fuzzy-ART. Os número de neurônios criados ao fim da etapa de treinamento da rede corresponde ao número de agrupamentos e cada caso é incluído ao grupo cujo vetor de pesos mais se assemelha a suas variáveis.

3.4 Calculando uma solução para um novo problema de acordo com um agrupamento

Para uma base de casos C , medidas de similaridade S_j escolhidas para cada variável v_j , medidas de similaridades globais S_{in} e S_{out} e um hipergrafo $H = (C, E)$ compatível com $R = F_{\emptyset,\alpha}$ para um operador residual \emptyset e um valor $\alpha \in (0,1]$, a questão é como calcular uma solução o^* apropriada para um novo problema p^* .

Esta solução é calculada a partir dos casos contidos no agrupamento onde os problemas descritos são de alguma maneira similares a p^* , denotado por $E^* = \{h \in E \mid \forall c_i = (p_i, o_i) \in h, S_{in}(p_i, p^*) > 0\}$. Para cada $h = \{c_1, c_2, \dots, c_r\} \in E^*$, é calculada a solução para p^* , denotada por o_h^* , usando uma função de agregação apropriada que leva em conta tanto o conjunto de soluções o_i quanto a similaridade entre cada p_i e p^* considerando os casos (p_i, o_i) em h . Por exemplo, se a função de

agregação é a média ponderada e as similaridades são agregadas usando S_{in} temos

$$o_h^* = \frac{\sum_{i=1,r} S_{in}(p_i, p^*) \times o_i}{\sum_{i=1,r} S_{in}(p_i, p^*)}$$

3.5 Determinando a força de um agrupamento em relação a um novo problema

Seja O^* o conjunto de soluções para p^* para o agrupamento E^* . Para selecionar a solução final o^* de O^* , pode-se agregar as soluções produzidas por um agrupamento ou pode-se simplesmente assumir como solução o agrupamento mais fortemente relacionado com p^* .

[11] propôs a *força* de um agrupamento $h = \{c_1, c_2, \dots, c_r\}, c_i = (p_i, o_i)$, em relação ao problema p^* como $str_f(h, p^*) = f(S_{in}(p_1, p^*), \dots, S_{in}(p_n, p^*))$ onde f é uma função de agregação adequada, como a média aritmética, uma *t-norma* ou uma *t-conorma*, por exemplo.

Neste trabalho, propomos que, além desta força de um agrupamento, fosse utilizado o fator de coesão de um agrupamento. Neste caso, a força do agrupamento é multiplicada por um fator de coesão que pode ser a semelhança de entrada média entre todos os casos que compõem o agrupamento.

4. Estudo de caso e resultados

Foram utilizados dados cedidos pela Secretaria do Estado de Minas Gerais que correspondem à prevalência de esquistossomose no estado apresentados originalmente em [7]. Minas Gerais possui 853 municípios sendo que a prevalência é conhecida para 197 deles (Figura 1) e o objetivo do trabalho é estimar a prevalência para os demais municípios.

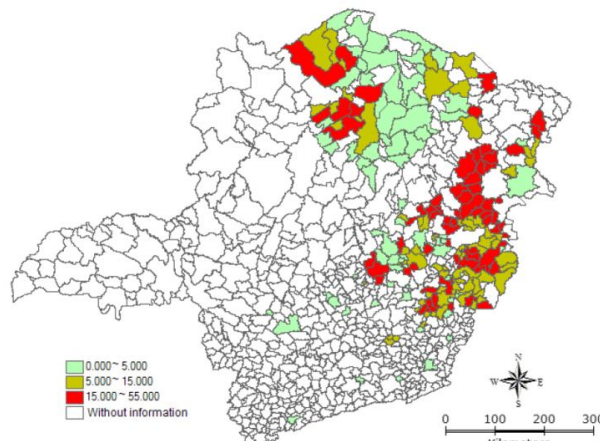


Figura 1- municípios mineiros cuja prevalência de esquistossomose é conhecida . Fonte: (Martins, 2008)

4.1 Experimentos originais

Nos experimentos originais, 86 variáveis independentes de vários tipos foram utilizados para calcular a prevalência. Variáveis de sensoriamento remoto (22), climáticas (6), socioeconômicas (34) e características de vizinhança (24). As variáveis de sensoriamento remoto são provenientes dos sensores MODIS (Moderate Resolution Imaging Spectroradiometer) e SRTM (Shuttle Radar Topography Mission), e supostamente relacionado ao tipo de habitat do caracol. As variáveis climáticas foram obtidas a partir da Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (INPE) e reflete as condições de sobrevivência de caramujos e as várias formas de larvas do *Schistosoma mansoni*. As variáveis socioeconômicas foram obtidas a partir do SNIU (Sistema Nacional de Indicadores Urbanos), tais como acesso a água potável e saneamento básico. As variáveis de características de vizinhança medem a disparidade entre os municípios vizinhos com relação as variáveis de renda, acesso a esgoto, educação, água e acúmulo de água.

Das 86 variáveis originais, um conjunto menor foi selecionado de acordo com testes utilizando regressão linear múltipla [7]; as variáveis independentes escolhidas foram aquelas que apresentaram alta correlação com a variável dependente e baixa correlação com as outras variáveis. Duas abordagens principais foram utilizadas: i) uma global, em que todos os municípios com prevalência da doença conhecida foram utilizados, tanto para construção ou validação de um modelo de regressão linear e, ii) uma regional, em que o estado foi dividido em quatro regiões homogêneas e um modelo de regressão linear foi criado para cada um deles. O número de variáveis independentes utilizadas nos experimentos variou; na abordagem global 5 variáveis foram utilizadas enquanto que na abordagem regional 2 variáveis foram utilizadas para a região R1, 5 para a região R2, 4 para a região R3 e 3 para a região R4. Em ambas as abordagens, global e regional, aproximadamente 2/3 das amostras foram utilizadas como conjunto de treinamento e 1/3 foi utilizado como conjunto de teste. O algoritmo SKATER foi utilizado para obter as regiões homogêneas no modelo regional; este algoritmo cria regiões tais que as áreas vizinhas com características semelhantes pertencem a mesma região (Figura 2).

Neste trabalho, utilizou-se como parâmetro de comparação apenas a região R3 pois esta é uma região mais completa levando-se em consideração a proporção de casos representantes de cada uma das classes.

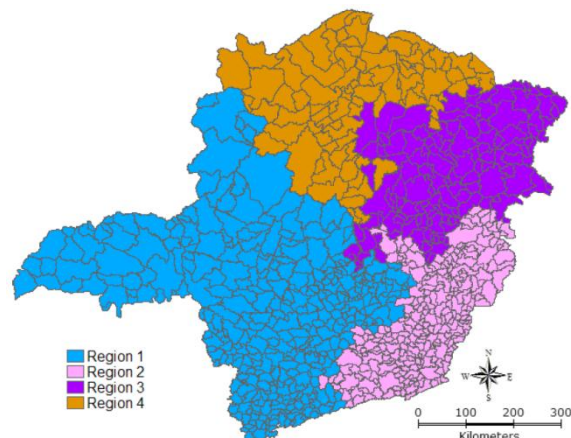


Figura 2- Regionalização obtida através do algoritmo SKATER Fonte: (Martins, 2008)

As prevalências foram classificadas como baixa ([0,5]%), média ([5,15]%) e alta ([15,100]%). A Tabela 2 reproduz os resultados de [8] com a precisão dos resultados (porcentual de casos corretamente classificados) obtidos por meio de regressão e abordagem da relação de similaridade ponderada para ambas as abordagens (regional e global).

Tabela 2 - Classificação com os modos de aprendizado R-Reg (Base regional e regressão), G-Reg (Base global e regressão), R-Sim (Base regional e similaridade) e G-Sim (Base global e similaridade)

Treinamento	R-Reg	G-Reg	R-Sim	G-Sim
R3(44)	72%	48%	62%	71%

A Tabela 3 reproduz os resultados de [8] com os resultados obtidos com árvore de decisão e regressão para todas as regiões, tanto para os dados de treinamento quanto para os dados de validação. Nota-se que para a região R1 a quantidade total de amostras era muito pequena (16), nas abordagens descritas não houve amostras de teste.

Podemos ver que, apesar de os resultados de treinamento para algumas regiões serem muito bons, os resultados do teste não são satisfatórios, com exceção da região R4.

Tabela 3 - Classificação, para conjuntos de treinamento e teste, com aprendizado: R-Reg (base regional e regressão), G-Reg (base global e regressão) e G-DT (base global e árvore de decisão).

Treinamento	R-Reg	G-Reg	G-DT
R3(44)	54,55% (24)	59,09% (26)	77,27% (34)
Teste	R-Reg	G-Reg	G-DT
R3(44)	28,57% (4)	42,86% (6)	35,71% (5)

Em [11], foram realizados experimentos com os mesmos dados utilizados em [7]. Agrupamentos foram usados para calcular o vetor de pesos e/ou calcular os resultados. Foram realizados 6 tipos de experimentos para calcular as prevalências. As estratégias são denotadas como segue:

- W-R : sem pesos e resultados calculados utilizando-se um único agrupamento;
- W-R+ : sem pesos e resultados calculados utilizando-se com vários agrupamentos;
- WR : com pesos e resultados calculados considerando-se um único agrupamento;
- WR+ : com pesos e resultados calculados para vários agrupamentos;

- W+R+: com pesos aprendidos sobre vários agrupamentos e saída usando o próprio agrupamento como base de treinamento;
- W++R+: similaridade ponderada usando agrupamentos expandidos e resultados para vários agrupamentos;

Os experimentos realizados em [11] utilizaram a função de agregação Gödel com corte de nível 1 e os agrupamentos foram calculados utilizando hipergrafos. Os resultados obtidos são apresentados na tabela 4.

Tabela 4 – Classificação para as abordagens regional e global para os experimentos W-R, W-R+, WR, WR+, W+R+, W++R+

Experimentos	W-R	W-R+	WR	WR+	W+R+	W++R+
R3	42,86% (6)	57,14% (8)	50 % (7)	71,42% (10)	14,29% (2)	57,14% (8)

Para este trabalho, primeiramente, fez-se um estudo a respeito de como os cortes de nível menores que 1 influenciariam na criação de agrupamentos. O uso de cortes de nível menores que 1 geraram um aumento no número de agrupamentos. Este aumento pode ser benéfico para a classificação, porém, para as relações de agregação utilizadas este aumento foi excessivo fazendo com que o processo de classificação ficasse inviável o cálculo de pesos para tantos agrupamentos.

Para as mesmas relações utilizadas em [11] utilizou-se também o fator de coesão entre os agrupamentos, aqui proposto. Os resultados obtidos são expostos na Tabela 5.

Tabela 5 – Classificação para as abordagens regional e global para os experimentos W-R-, W-R+, W+R-, W+R+, W++E-R+, W++E+R+ utilizando fator de coesão

Experimentos	W-R	W-R+	WR	WR+	W+R+	W++R+
R3	42,86% (6)	57,14% (8)	42,86 % (6)	57,14% (8)	64,28% (9)	57,14% (8)

Utilizando-se a RNA Fuzzy-ART observou-se que os melhores resultados foram obtidos quando a rede foi treinada em 600 iterações e utilizou-se o limiar de vigilância igual a 0,45. Neste trabalho, considerou-se que um bom número de agrupamentos seria algo em torno de 3 agrupamentos. Este valor foi estipulado porque a base de casos utilizada como teste é composta de 3 classes. Outro fato que deve ser observado é que quando utilizamos a rede neural não acontece de um determinado caso pertencer a mais de um agrupamento.

Foram realizados 3 treinamentos para a RNA. No primeiro, a rede foi treinada com os atributos que formavam cada um dos casos. Depois, efetuou-se o treinamento a partir da relação de similaridade S_{in} , considerando-se que os casos deveriam ser agrupados com base somente em sua similaridade de entrada. Por último, foi usada a relação $F_{\theta}(c_a, c_b)$ que efetua uma combinação entre as similaridades de entrada e de saída. O número de agrupamentos para cada um dos treinamentos foi 4, 5 e 3 respectivamente. A tabela 6 mostra os resultados obtidos quando os agrupamentos foram calculados usando o rede Fuzzy-ART sem o fator de coesão entre os agrupamentos gerados.

Tabela 6 – Resultados obtidos quando os agrupamentos foram calculados utilizando a rede fuzzy-ART

Experimentos	W-R	W-R+	WR	WR+	W+R+	W++R+
	42,86% (6)	42,86% (6)	42,86% (6)	64,29% (9)	42,86% (6)	42,86% (6)
	42,86% (6)	64,29% (9)	42,86% (6)	35,71% (5)	50% (7)	42,86% (6)
	42,86% (6)	57,14% (8)	42,86% (6)	21,43% (3)	35,71% (5)	42,86% (6)

5. Conclusão

Este trabalho concentra-se na tarefa de classificação de dados. Até o presente momento, comparou-se os resultados até aqui obtidos com várias estratégias que utilizam uma base de dados criada para estimar a prevalência da esquistossomose para uma região do Brasil [7]. Os resultados obtidos foram em alguns casos melhores que os apresentados em [7] e [8] para lidar com o mesmo problema, o que mostra que a abordagem apresentada é promissora.

Através dos experimentos realizados neste trabalho observou-se que em alguns casos a aplicar cortes de nível na função de agregação pode gerar um número excessivo de agrupamentos. O que torna a técnica inviável, pois a etapa mais custosa deste processo é o cálculo dos vetores de pesos de cada um dos casos.

Tal dificuldade pode ser contornada utilizando-se a rede neural artificial Fuzzy-ART para calcular os agrupamentos. Porém, sua utilização é mais indicada quando a base de casos é previamente conhecida, pois o processo de treinamento pode ser mais custoso e gerar agrupamentos desnecessários caso seus parâmetros sejam mal ajustados. Observou-se também que os resultados obtidos utilizando esta abordagem não foram tão bons quanto os obtidos quando hipergrafos foram utilizados pra se calcular os agrupamentos.

Outro ponto a se observar é que o fator de coesão utilizado melhora significativamente os resultados quando utilizamos os agrupamentos calculados tanto para cálculo dos pesos quanto para o cálculo dos resultados finais da classificação.

Referências

1. Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. 7(1):39--59, 1994.
2. Abel, M.. Um estudo sobre raciocínio baseado em casos. UFRS, 1996.
3. Armengol, E.; Esteva, F.; Godo, L.; Torra, V.. On learning similarity relations in fuzzy case-based reasoning. *Trans. on Rough Sets*, pp 14--32, 2004.
4. Dubois, D.; Esteva, F.; Garcia, P.; Godo, L.; Mantaras, R. L.; Prade, H. Fuzzy set modelling in case-based reasoning. *International Journal of Intelligent Systems*, John Wiley & Sons, Ltda., v. 13, n. 4, p. 345--373, April 1998. 38, 39

5. Fanoiki, T.; Drummond, I.; Sandri, S.. Case-based reasoning retrieval and reuse using case resemblance hypergraphs. Proc IEEE Int. Conf Fuzzy Systems (FUZZ'10), 2010.
6. Kolodner, J. *Cased-based reasoning*. Morgan Kaufmann, San Mateo, CA, 1993.
7. Martins, F.T.; Freitas, C.; Dutra, L.; Sandri, S.; Drummond, I.; Fonseca, F.; Guimarães, R.; Amaral, R.; Carvalho, O.. Risk mapping of Schistosomiasis in the state of Minas Gerais, Brazil, using MODIS and socioeconomic spatial data. IEEE Trans. on Geoscience and Remote Sensing, v47, n11, pp 3899-3908, 2008.
8. Martins-Bedê, F.T.; Godo, L.; Sandri, S. A.; Dutra, L.V.; Freitas, C.C.; Carvalho, O.S.; Guimaraes, R.J.; Amaral, R.S.. Classification of schistosomiasis prevalence using fuzzy case-based reasoning. International Work-Conference on Artificial Neural Networks (IWANN'09), v5517, pp 1053--1060. Springer, 2009.
9. Grossberg et al. Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptative Resonance System. *Neural Networks*, Vol. 4, pp. 759-771, 1991.
10. Ruspini, E.; Bonissone, P.; Pedrycz, W. Handbook of fuzzy computation. London: IOP Publishing Ltd, 1998. 504 p. ISBN 9780750304276.
11. Sandri, S., Mendonça, J. H., Martins-Bedê, F.T. Weighted Fuzzy Similarity Relations Case-Based Reasoning: a case study in classification. World Conference Computational Intelligence, 2012a.
12. Sandri, S., Lesot, M. J. A Fuzzy Residuated Approach to Case-based Reasoning. IPMU, Itália, 2012.
13. Silva, N. C. Utilização de operadores genéticos para otimizar classificadores neurais não-supervisionados de imagens. Brasília. 200p. Tese de Doutorado em Geociências - Universidade de Brasília, 2002.
14. Torra, V.. On the learning of weights in some aggregation operators: the weighted mean and OWA operators. *Math. and Soft Comp.* 6, 2000.
15. Vitorino, T. A. S., Caminhas, W. M.. Raciocínio baseado em casos: conceitos e aplicações. Dissertação de mestrado. Universidade Federal de Minas Gerais. Brasil, 2009.
16. Zadeh, L. A. *Fuzzy Sets*, Information and Control, 8: 338-353., 1965
17. Zadeh, L. Similarity relations and fuzzy orderings. *Information Sciences*, Elsevier Science Ltd., v. 3, n. 2, p. 177–200, April 1971.