

Raciocínio baseado em casos: uma abordagem utilizando relações de similaridade difusa

Jonas Henrique Mendonça¹, Sandra Sandri², Flávia Martins-Bedê¹, Ricardo Guimarães³,
Omar Carvalho⁴

¹Programa de Mestrado ou Doutorado em Computação Aplicada – CAP
Instituto Nacional de Pesquisas Espaciais – INPE

²Laboratório Associado de Computação e Matemática Aplicada – LAC
Instituto Nacional de Pesquisas Espaciais – INPE

³Laboratório de Geoprocessamento / Instituto Evandro Chagas / SVM / MS

⁴Centro de Pesquisas René Rachou / FIOCRUZ

jonas.henrique01@gmail.com, sandri@lac.inpe.br, toledomartins@gmail.com, ricardojpsg@gmail.com,
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 and groupings that would later be used in data classification tasks.*

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 obtained for this case are based, in most cases, superior to approaches previously applied.

Resumo. *Este trabalho é um estudo de uma abordagem de raciocínio baseado em casos utilizando relações de similaridade difusa ponderada. Foram utilizadas diversas métricas para obtenção dos vetores de pesos e agrupamentos que seriam posteriormente utilizados em tarefas de classificação de dados.*

Utilizamos como estudo de caso um problema do mundo real para estimar a prevalência da esquistossomose em uma região do Brasil. Os resultados obtidos para esta base de casos são, na maioria dos casos, superiores a abordagens anteriormente aplicadas.

Palavras-chave: *Lógica difusa, relações de similaridade, agrupamento*

1. Introdução

A técnica de reconhecimento baseada em casos (RBC) (kolodner, 1993) se propõe a resolver um problema usando um princípio que pode ser declarado como “problemas semelhantes tem soluções semelhantes” (Aamodt et al, 1994). O primeiro passo deste procedimento consiste em recuperar problemas na base de casos que são semelhantes ao problema considerado, ou seja, determinar os casos que são relevantes para a solução do problema. A segunda etapa consiste em reutilizar as soluções desses problemas, adaptando-os para o problema considerado.

Pesos podem ser anexados aos casos, de maneira que alguns destes possam ser considerados mais importantes para uma determinada aplicação. 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.

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 Fanoiki et. al. (2010) e, outra maneira para se calcular estes fragmentos é a utilização de redes neurais artificiais (RNAs). 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 (Sandri, 2012b). 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 Sandri et. al. (2012a) 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 acerca desta metodologia para raciocínio baseado em casos aplicado a uma base de casos. Este trabalho está organizado da seguinte maneira: a seção 2 mostra as definições básicas deste trabalho. A metodologia utilizada é apresentada na seção 3. Apresentamos um estudo de caso e os resultados obtidos na seção 4 e finalmente, a seção 5 traz a conclusão deste trabalho.

2. Fundamentação teórica

Nesta seção apresentamos as definições e notações utilizadas neste trabalho.

2.1 Lógica difusa e relações de similaridade

A teoria dos conjuntos difusos foi desenvolvida a partir de 1965 por Lotfi Zadeh (Zadeh, 1965). Formalmente, um conjunto nebuloso C do universo de discurso A é definido por uma função de pertinência $\mu(x): A \rightarrow [0,1]$. Essa função associa a cada elemento de x de A o grau $\mu(x)$ com o qual pertence a C . A função de pertinência $\mu(x)$ indica o grau de compatibilidade entre x e o conceito expresso por C .

Na lógica difusa 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*.

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 (1971) como uma generalização da definição clássica de relações de equivalência.

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 Silva (2002), 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 difusos.

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 (Abel, 1996). De acordo com (Aamtod et. al., 1994), 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 Vistorino et. al.(2009), 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 Sandri et. al. (2012a). 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:

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 3.1.

$$F_\emptyset(c_a, c_b) = \begin{cases} 0, & \text{se } S_{in}(p_a, p_b) = 0 \\ \emptyset(S_{in}(p_a, p_b), S_{out}(o_a, o_b)), & \text{senão} \end{cases} \quad (3.1)$$

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 3.1 pode-se usar relações como Goguen e seus respectivos cortes de nível como mostrado na equação 3.2.

$$\forall \alpha \in (0,1], F_{\emptyset, \alpha}(c_i, c_j) = \begin{cases} 1, & \text{se } F_\emptyset(c_i, c_j) \geq \alpha \\ 0, & \text{caso contrário} \end{cases} \quad (3.2)$$

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:

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

Na segunda, utiliza-se a RNA Fuzzy-ART. O 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 temos 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 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 $o_h^* = \sum_{i=1,r} \frac{S_{in}(p_i, p^*) \times o_i}{\sum_{i=1,r} S_{in}(p_i, p^*)}$.

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^* .

Sandri et. al.(2012a) 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 (Martins et. al., 2008). 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.

4.1 Experimentos originais

Nos experimentos originais, 86 variáveis independentes de vários tipos foram utilizados para calcular a prevalência. Das 86 variáveis originais, um conjunto menor foi selecionado de acordo com testes utilizando regressão linear múltipla (Martins et. al., 2008); as variáveis independentes escolhidas foram aquelas que apresentaram alta correlação com a variável dependente e baixa correlação com as outras variáveis. Foi utilizada uma abordagem regional, em que o estado foi dividido em quatro regiões homogêneas e um modelo de regressão linear foi criado pra cada um deles. O algoritmo SNAKER 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).

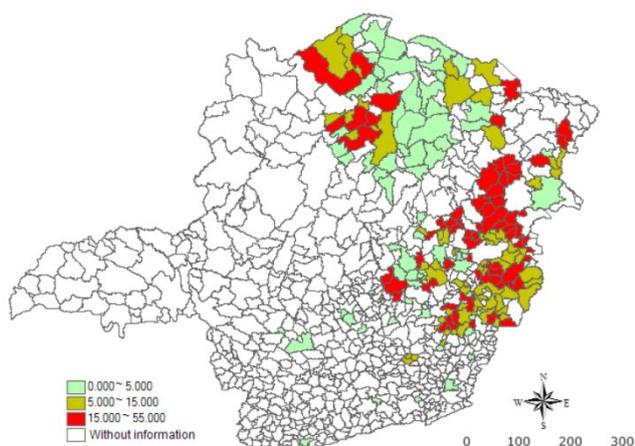


Figura 2- municípios mineiros cuja prevalência de esquistossomose é conhecida. Fonte: (Martins, 2008)

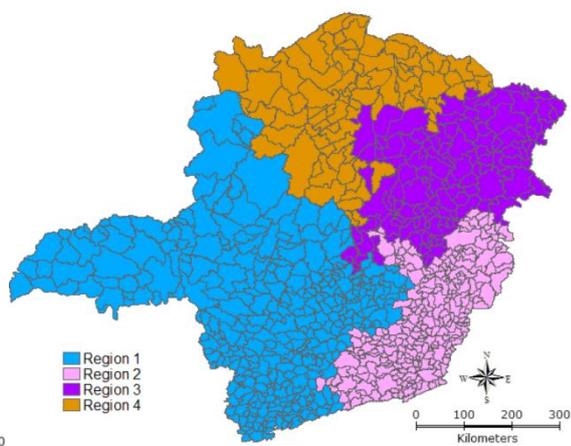


Figura 1- Regionalização obtida através do algoritmo SNAKER Fonte: (Martins, 2008)

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.

As prevalências foram classificadas como baixa ($[0,5\%)$), média ($[5,15\%)$) e alta ($[15,100\%]$). A Tabela 4.1 reproduz os resultados de (Martins-Bedê, 2009) com a precisão dos resultados (porcentual de casos corretamente classificados) obtidos por meio de regressão e abordagem da relação de similaridade ponderada.

Tabela 4.1 - 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
	72%	48%	62%	71%

A Tabela 4.2 reproduz os resultados de (Martins-Bedê, 2009) com os resultados obtidos com árvore de decisão e regressão, tanto para os dados de treinamento quanto para os dados de validação.

Tabela 4.2 - 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%	59,09%	77,27%
Teste	R-Reg	G-Reg	G-DT
R3(14)	28,57%	42,86%	35,71%

4.2 Resultados obtidos

Em Sandri et. al. (2012a) foram realizados experimentos com os mesmos dados utilizados em Martins et. al. (2008). Agrupamentos foram usados para calcular o vetor de pesos e/ou calcular os resultados. Foram realizados seis tipos de experimentos para calcular os resultados. 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 saídos usando o próprio agrupamento como base de treinamento;

W++R+: similaridade ponderada usando agrupamentos expandidos e resultados para vários agrupamentos;

Os experimentos realizados em Sandri et. al. (2012a) 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 4.3.

Tabela 4.3 – 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%	57,14%	50 %	71,42%	14,29%	57,14%

Neste trabalho, para as mesmas relações utilizadas em Sandri et. al. (2012a) utilizou-se também o fator de coesão entre os agrupamentos. Os resultados obtidos são expostos na Tabela 4.4.

Tabela 4.4 – 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%	57,14%	42,86 %	57,14%	64,28%	57,14%

A RNA Fuzzy-ART também foi utilizada para calcular os agrupamentos de casos. Realizou-se três tipos de treinamento 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 $F_{\phi}(c_a, c_b)$ 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 4.5 mostra os resultados obtidos quando os agrupamentos foram calculados usando o rede Fuzzy-ART sem o fator de coesão entre os agrupamentos gerados.

Tabela 4.5 – Resultados obtidos quando os agrupamentos foram calculados utilizando a rede fuzzy-ART

Experimentos	W-R	W-R+	WR	WR+	W+R+	W++R+
T1	42,86%	42,86%	42,86%	64,29%	42,86%	42,86%
T2	42,86%	64,29%	42,86%	35,71%	50%	42,86%
T3	42,86%	57,14%	42,86%	21,43%	35,71%	42,86%

Para efeito de comparação utilizou-se a RNA MLP para classificação destes casos obtendo-se uma taxa de acerto de 64,29%.

5. Conclusão

Este trabalho concentra-se na tarefa de classificação de dados. Até o presente momento, compararam-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. Os resultados obtidos foram em alguns casos melhores que os apresentados na literatura e os apresentados pela RNA MLP, que é uma técnica muito utilizada para classificação de dados. Isto mostra que a abordagem apresentada é promissora.

O uso da RNA Fuzzy-ART gerou agrupamentos contendo casos que, originalmente, pertencem a classes diferentes. Tal efeito faz com que os treinamentos utilizando toda a base de casos gere resultados mais significativos que os obtidos a partir da ponderação entre os casos do agrupamento. Porém, em nenhum dos testes realizados, os resultados foram superiores aos obtidos quando foram calculados os agrupamentos a partir de hipergrafos.

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

- Aamodt and E. Plaza. Case-based reasoning: Foundational issues, methodological variations, and system approaches. 7(1):39--59, 1994.
- Abel, M.. Um estudo sobre raciocínio baseado em casos. UFRS, 1996.
- Beale, R. & Jackson, T. *Neural Computing: An Introduction. Institute of Physics Publishing.* 1990.
- 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
- 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.
- Kolodner, J. *Cased-based reasoning*. Morgan Kaufmann, San Mateo, CA, 1993.
- 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.
- 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.
- Grossberg et al. Fuzzy ART: Fast Stable Learning and Categorization of Analog Patterns by an Adaptive Resonance System. *Neural Networks*, Vol. 4, pp. 759-771, 1991.
- Ruspini, E.; Bonissone, P.; Pedrycz, W. *Handbook of fuzzy computation*. London: IOP Publishing Ltd, 1998. 504 p. ISBN 9780750304276.
- 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.
- Sandri, S., Lesot, M. J. A Fuzzy Residuated Approach to Case-based Reasoning. *IPMU*, Itália, 2012.
- 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.
- Torra, V.. On the learning of weights in some aggregation operators: the weighted mean and OWA operators. *Math. and Soft Comp.* 6, 2000.
- Vitorino, T. A. S., Caminhas, W. M.. *Raciocínio baseado em casos: conceitos e aplicações*. Dissertação de mestrado. Universidade Federal de Minas Gerais. Brasil, 2009.
- Zadeh, L. A. *Fuzzy Sets*, *Information and Control*, 8: 338-353., 1965
- Zadeh, L. Similarity relations and fuzzy orderings. *Information Sciences*, Elsevier Science Ltd., v. 3, n. 2, p. 177–200, April 1971.