

Um Algoritmo para Selecionar Regras de Conhecimento Utilizando Medidas de Avaliação de Regras *

Flávia Cristina Bernardini ¹ Maria Carolina Monard ¹ Huei D. Lee ^{2 1}
Sandro C. Esteves ³

¹ Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação
Caixa Postal 668, 13560-970 - São Carlos, SP, Brasil
e-mail: {fbernard,mcmonard}@icmc.usp.br

² Universidade Estadual do Oeste do Paraná
LABI - Laboratório de Bioinformática
Caixa Postal 961, 85870-900 - Foz do Iguaçu, PR, Brasil
huei@unioeste.br

³ Centro de Referência em Infertilidade Masculina — Androfert
Av. Dr. Heitor Pentead, 1464, Taquaral, 13075-460 Campinas, SP
sandro.esteves@androfert.com.br

Resumo: Em aprendizado de máquina supervisionado, não somente a precisão do classificador deve ser considerada mas também a análise do conhecimento induzido. No caso de aprendizado simbólico a análise das regras que constituem o classificador é uma atividade muito importante. O objetivo deste trabalho é propor um algoritmo de seleção de regras de conhecimento, denominado \mathcal{RCO} , que utiliza como entrada um conjunto de classificadores simbólicos e uma medida de avaliação para selecionar regras. O \mathcal{RCO} fornece como resultado o melhor conjunto de regras segundo a medida de seleção utilizada. Assim, o conjunto construído pelo \mathcal{RCO} permite ao usuário se concentrar em um conjunto de regras menor que o conjunto de regras original, o qual foi fornecido ao \mathcal{RCO} como entrada, e com propriedades específicas. A utilização do \mathcal{RCO} é ilustrada em um estudo de caso do mundo real.

Palavras Chaves: aprendizado de máquina simbólico, medidas de avaliação de regras, seleção de regras de conhecimento.

1 Introdução

Em Aprendizado de Máquina — AM — supervisionado, além da indução de classificadores precisos, é também importante analisar o conhecimento induzido com o objetivo de descobrir novo conhecimento. Nesse caso, há um forte apelo para usar algoritmos de AM simbólicos, nos quais o modelo induzido pode ser diretamente interpretado pelo usuário/especialista do domínio.

Porém, dado um conjunto de casos (exemplos) rotulados, nem sempre é possível induzir um classificador suficientemente preciso, ou seja, que classifique novos casos com a precisão exigida pelo especialista. Ainda assim, é possível selecionar conjuntos de regras que fazem parte desses classificadores insuficientemente precisos mas que cobrem bem alguns desses casos. Domínios nos quais são induzidos tais tipos de classificadores são freqüentemente encontrados na área médica. Uma situação típica ocorre quando existe mais de um tipo de análise laboratorial que

*Trabalho realizado com o auxílio da FAPESP, Brasil, Proc. N° 02/06914-5.

auxilie o médico no diagnóstico de alguma doença. Muitas vezes, a análise laboratorial mais precisa é mais custosa, ou em termos econômicos ou por ser desconfortável para o paciente. Tipicamente, para analisar essas situações, é necessário ter informações de um conjunto de pacientes (casos) para os quais ambas análises foram realizadas. O rótulo de cada caso é fornecido pelo resultado da análise mais precisa (custosa), enquanto que os atributos são fornecidos pela análise menos custosa.

Neste trabalho é apresentado o algoritmo *RCO* (*Rule Combination*), o qual pode ser utilizado para fornecer subsídios ao especialista na solução parcial desse tipo de problema, ou seja, construir conjuntos de regras, segundo alguma visão de seleção de regra, que cubram bem alguma região do domínio. O *RCO* faz parte de um dos módulos de um sistema denominado *RuleSystem*, detalhadamente descrito em (Bernardini 2002; Gomes 2002), o qual foi desenvolvido na linguagem de programação lógica Prolog. O principal objetivo do *RCO* é selecionar regras induzidas por algoritmos de AM simbólico, utilizando diversas medidas de avaliação de regras de conhecimento. Os conjuntos de regras construídos pelo *RCO* segundo essas diferentes visões (medidas) podem ser analisados e comparados pelo especialista, auxiliando-o, assim, em possíveis tomadas de decisão ou no descobrimento de novo conhecimento relacionado ao processamento de sêmen humano. O conjunto de dados é composto por casos nos quais os atributos são dados por um exame mais simples, denominado espermograma, enquanto os rótulos são dados por um exame mais custoso, denominado processamento de sêmen.

Este trabalho está organizado em 5 seções. Na Seção 2 são descritos alguns conceitos de AM e as medidas de avaliação de regras utilizadas no estudo de caso; o algoritmo *RCO* proposto encontra-se na Seção 3 e o estudo de caso realizado, na Seção 4. Finalmente, na Seção 5 são apresentadas as conclusões do trabalho.

2 Aprendizado de Máquina Simbólico e Avaliação de Regras

No problema padrão de AM supervisionado, ao algoritmo de aprendizado é dado um conjunto de exemplos de treinamento $S = \{(\mathbf{x}_i, y_i)\}, i = 1, \dots, M$, onde $\mathbf{x}_i = (x_1, \dots, x_N)$, com $x_j, j = 1, \dots, N$ atributos discretos ou contínuos, e y_i o atributo classe. No caso de classificação, tratado neste trabalho, o atributo classe y_i é discreto, ou seja, $y_i \in \{C_1, C_2, \dots, C_{NCL}\}$. A partir do conjunto S , um *classificador* \mathbf{h} é induzido. No caso de aprendizado simbólico, \mathbf{h} pode ser constituído de um conjunto de regras *if-then* $\{R_1, R_2, \dots, R_{NR}\}$, *i.e.*, regras do tipo *Corpo* \rightarrow *Cabeça* ou *Body* \rightarrow *Head* ou, simplesmente, $B \rightarrow H$. Nesse caso, dado um exemplo \mathbf{x} e uma regra R , se o *Corpo* de R é verdade para esse exemplo, então R cobre \mathbf{x} . É possível avaliar cada uma das regras que constituem \mathbf{h} separadamente. Para avaliar regras, várias medidas de avaliação de regras têm sido propostas na literatura (Freitas 1998b; Freitas 1998a; Hilderman & Hamilton 1999). Neste trabalho, são utilizadas medidas de avaliação de regras unificadas no *framework* proposto por (Lavrac, Flach, & Zupan 1999). Para as medidas nesse *framework*, é necessário um conjunto de exemplos S e algumas informações a respeito de cada regra $R_u \in \mathbf{h}$: o número de exemplos em S para os quais H é verdade e B é verdade (hb); o número de exemplos em S para os quais H é falso e B é verdade ($\bar{h}b$); o número de exemplos em S para os quais H é falso e B é falso ($\bar{h}\bar{b}$); e o número de exemplos em S para os quais H é verdade e B é falso ($h\bar{b}$). Após determinar tais informações para cada uma das regras R_u , os valores encontrados podem ser expressos sob a forma de frequências relativas, as quais são utilizadas como uma estimativa de probabilidade (Prati, Baranauskas, & Monard 2001). Por exemplo, considerando a informação hb ,

Tabela 1: Medidas de Avaliação de Regras Utilizadas

Precisão de uma regra:
$Acc(B \rightarrow H) = P(H B) = \frac{hb}{b}$
Sensitividade de uma regra:
$Sens(B \rightarrow H) = P(B H) = \frac{hb}{h}$
Especificidade de uma regra:
$Spec(B \rightarrow H) = P(\bar{B} \bar{H}) = \frac{\bar{h}\bar{b}}{\bar{h}}$
Cobertura de uma regra:
$Cov(B \rightarrow H) = P(B) = \frac{b}{N}$
Suporte de uma regra:
$Sup(B \rightarrow H) = P(HB) = \frac{hb}{N}$
Satisfação de uma regra:
$Sat(B \rightarrow H) = \frac{(P(\bar{H}) - P(\bar{H} B))}{P(\bar{H})} = \frac{\frac{\bar{h}}{N} - \frac{\bar{h}\bar{b}}{b}}{\frac{\bar{h}}{N}} = 1 - \left(\frac{\bar{h}\bar{b}}{b} \cdot \frac{N}{\bar{h}}\right)$

pode-se dizer que $P(HB) = f_{bh} = \frac{hb}{N}$. De forma semelhante podem ser determinados os valores das probabilidades $P(\overline{HB})$, $P(\overline{HB})$ e $P(H\overline{B})$. Conhecidas essas probabilidades, os valores de $P(B)$, $P(\overline{B})$, $P(H)$ e $P(\overline{H})$ podem ser determinados. Por exemplo: $P(B) = P(\overline{HB}) + P(HB)$ ou $P(\overline{B}) = P(H\overline{B}) + P(\overline{HB})$. Assim, nesse *framework*, a precisão de uma regra $B \rightarrow H$, que mede o quanto uma regra é específica para o problema, é definida como a probabilidade condicional de H ser verdade dado que B é verdade, ou seja, $Acc(B \rightarrow H) = P(H|B) = \frac{P(HB)}{P(B)} = \frac{hb}{b}$. De forma semelhante são definidas as outras medidas de regras. As medidas utilizadas no estudo de caso realizado encontram-se na Tabela 1.

3 O RCO

Dado um conjunto de hipóteses (classificadores) $\mathcal{H} = \{\mathbf{h}_1, \dots, \mathbf{h}_L\}$, no qual cada hipótese \mathbf{h}_i é expressa como um conjunto de regras, ou seja, $\mathbf{h}_i = \{R_{i_1}, R_{i_2}, \dots, R_{i_{NR}}\}$; um conjunto de exemplos de treinamento S ; e uma medida de avaliação de regras *Measure*, o objetivo do algoritmo RCO (*Rule Combination*) — Algoritmo 1 — é encontrar um conjunto de regras *BestRules* com os melhores valores da medida *Measure* que cobrem esses exemplos. O algoritmo segue basicamente os seguintes passos:

1. Verifica se o conjunto de exemplos S está vazio. Se sim, a iteração pára;
2. Calcula *Measure* para todas as regras $R \in \mathcal{H}$, com exceção das regras *default*¹;
3. Retira a regra *Rule* com melhor valor *Measure* dentre as L hipóteses;
4. Retira do conjunto de exemplos S todos os exemplos que são cobertos pela regra *Rule*. Se nenhum exemplo for coberto, a iteração pára;
5. Adiciona a regra *Rule* ao conjunto de regras *BestRules*;
6. Volta ao passo inicial.

A seguir é apresentado o estudo de caso realizado utilizando o RCO.

4 Um Estudo de Caso — Processamento de Sêmen Diagnóstico

O estudo de caso considerado tem como objetivo tentar prever os resultados de um exame (análise) laboratorial denominado processamento de sêmen, mais custoso, baseado no resultado de um exame laboratorial menos custoso, denominado espermograma. Deve ser observado que o exame de processamento de sêmen pode chegar a ser aproximadamente 80% mais custoso financeiramente que o espermograma. O conjunto de dados é construído rotulando os resultados obtidos de espermogramas com os resultados obtidos de exames de processamento de sêmen, ambos realizados no mesmo paciente. A importância do processamento de sêmen para a reprodução assistida é descrita em (Lee 2000). Maiores detalhes sobre o estudo de caso aqui apresentado pode ser encontrado em (Bernardini & Monard 2002).

O exame de processamento de sêmen permite (a) quantificar com precisão a qualidade do sêmen (processamento de sêmen diagnóstico); e (b) recuperar a maior quantidade possível de espermatozoides para a utilização na reprodução assistida (processamento de sêmen terapêutico). A quantidade de espermatozoides recuperados pelo processamento de sêmen influencia na escolha da técnica que será utilizada no tratamento. São utilizadas três técnicas no tratamento para a reprodução assistida: (1) Inseminação Intra Uterina — IUI; (2) Fertilização In Vitro — FIV; e (3) Injeção Intracitoplasmática do Espermatozoide no Oócito — ICSI.

¹Uma regra *default* é caracterizada por possuir *Corpo* vazio e sua *Cabeça* é definida pela classe com maior distribuição nos exemplos do conjunto de exemplos inicial.

Algoritmo 1 \mathcal{RCO}

Require: $\mathcal{H} = \{\mathbf{h}_l, l = 1, \dots, L\}$: conjunto de hipóteses;

$S = \{(\mathbf{x}_i, y_i), i = 1, \dots, M\}$: conjunto de exemplos;

$Measure$: medida de regra.

1: **procedure** $rco(\mathcal{H}, S, Measure)$

2: $removeDefaultRules(\mathcal{H})$; {Remover todas as regras *default* de \mathcal{H} }

3: **loop**

4: **if** $vazio(S)$ or $vazio(\mathcal{H})$ **then**

5: **exit loop**; {Não há mais exemplos de treinamento ou não há mais regras em \mathcal{H} para processar}

6: **end if**

7: $calculateMeasuresOverAllSetOfRules(Measure)$; {Calcular a medida $Measure$ para todas as hipóteses em \mathcal{H} }

8: $Rule := \max(\mathcal{H}, Measure)$; {Encontrar a regra $Rule$ em \mathcal{H} com maior medida $Measure$ }

9: $S_{cov} := cov(S, Rule)$; {Criar o conjunto de exemplos cobertos pela regra $Rule$ }

10: **if** $vazio(S_{cov})$ **then**

11: **exit loop**; {Nenhum exemplo foi coberto pela regra $Rule$ }

12: **end if**

13: $S := S - S_{cov}$; {Retirar de S os exemplos contidos em S_{cov} }

14: $BestRules := BestRules \cup \{Rule\}$; $\mathcal{H} := \mathcal{H} - \{Rule\}$; {Colocar $Rule$ em $BestRules$ e retirar $Rule$ de \mathcal{H} }

15: **end loop**

16: **return** $BestRules$;

O processamento de sêmen permite que o maior número de espermatozoides móveis (motilidade) seja recuperado. Os espermatozoides são classificados, segundo sua motilidade, em graus A (maior grau de motilidade), B (menor grau de motilidade), C (espermatozoides que se movem em círculos) e D (espermatozoides que são imóveis). Os graus de motilidade são atributos medidos tanto no espermograma quanto no processamento de sêmen. Os casos que compõem esse conjunto de dados foram extraídos de casos reais do Centro de Referência em Infertilidade Masculina — Androfert² — em Campinas, SP, fornecidos pelo próprio especialista Dr. Sandro C. Esteves.

Tabela 2: Características do Conjunto de Dados

# Ex	#Conf	# At (con,nom)	Classe/%	EM	VD
240	1 (0.4%)	21 (10,11)	1 42.3% 2 22.1% 3 35.0%	57.1% em 1	S

4.1 Descrição do Conjunto de Dados

A Tabela 2 sumariza as características do conjunto de dados utilizado neste estudo. Ela mostra o número de exemplos (#Ex), número e porcentagem de exemplos com valores conflitantes (que possuem o mesmo atributo-valor mas têm diferentes classificações) (#Dup/Conf), número de atributos (# At) contínuos (con) e nominais (nom), distribuição das classes, o erro majoritário (EM) e se o conjunto de dados tem ao menos um valor desconhecido (VD).

4.2 Descrição do Experimento

O experimento conduzido tem por objetivo construir conjuntos de regras $BestRules$ utilizando o \mathcal{RCO} e algumas das medidas de qualidade de regras implementadas no \mathcal{RCO} . Inicialmente, utilizando o conjunto total de exemplos S , com 240 exemplos, foi induzida uma hipótese, a qual contém 33 regras e uma precisão baixa de aproximadamente 59%³ (ou erro alto de aproximadamente 41%). A taxa de precisão obtida na indução dessa hipótese mostra que esse é um caso típico onde não é possível induzir um classificador suficientemente preciso.

²<http://www.androfert.com.br/>

³Essa taxa de precisão bem como as taxas de precisão — e erro — apresentadas neste trabalho foram obtidas utilizando a técnica de 10-fold cross-validation. Em (Baranauskas & Monard 2000) existe uma explicação de como essa técnica funciona, dentre outras técnicas de estimativa de precisão de classificadores.

Nesse problema, então, foi pensado em utilizar o \mathcal{RCO} para construir conjuntos de regras com as melhores regras segundo visões distintas de medidas de avaliação de regras, constantes na Tabela 1. Deve ser observado que, para utilizar o \mathcal{RCO} , é interessante utilizar vários classificadores com regras distintas (não necessariamente classificadores totalmente distintos). Como é possível selecionar regras que cubram bem alguns exemplos ou alguma região do domínio, pensou-se em dividir o conjunto de exemplos inicial em subconjuntos a serem utilizados para induzir novos classificadores. Esses novos classificadores são utilizados como entrada para o \mathcal{RCO} . Assim, o conjunto S foi dividido aleatoriamente em quatro subconjuntos disjuntos S_1 , S_2 , S_3 e S_4 de exemplos com o mesmo número de exemplos em cada um deles — 60 exemplos. Utilizando os subconjuntos S_1 , S_2 e S_3 , foram induzidas 3(três) hipóteses: \mathbf{h}_1 , \mathbf{h}_2 e \mathbf{h}_3 , usando o algoritmo de aprendizado $\mathcal{CN}2$ (Clark & Niblett 1989). Os dados relacionados a todas as hipóteses induzidas constam na Tabela 3⁴: erro aparente do classificador (EA), estimativa de erro do classificador nas classes obtida utilizando 10-fold cross-validation (EC), estimativa de erro do classificador (Erro — $e \pm se^5$), obtida utilizando 10-fold cross-validation, e número de regras de cada classificador induzido.

Utilizando o conjunto de exemplos S_4 e as medidas de avaliação de regras constantes na Tabela 1, ou seja, precisão, cobertura, satisfação, sensibilidade, especificidade e suporte, foram construídos 6(seis) conjuntos $BestRules$. Esses conjuntos foram denominados, respectivamente, BR_{accR} , BR_{covR} , BR_{satR} , BR_{sensR} , BR_{specR} e BR_{supR} . Na Tabela 4 é mostrado o número de exemplos restantes, após a condição de parada do \mathcal{RCO} , na construção desses conjuntos de regras e o número de regras em cada conjunto.

4.3 Análise dos Resultados

Quanto ao número de regras em cada conjunto de regras $BestRules$ (Tabela 4), pode ser observado que esse número é geralmente menor que o número de regras que constituem cada um dos classificadores \mathbf{h}_1 , \mathbf{h}_2 e \mathbf{h}_3 . Quanto às regras selecionadas pelo \mathcal{RCO} para os 6(seis) critérios utilizados, foi observado que foram selecionadas um total de 17 regras diferentes que participam dos classificadores construídos pelo \mathcal{RCO} . Analisando os resultados foi observado que: (1) regras que pertencem a dois ou mais conjuntos $BestRules$ foram selecionadas em ordens distintas nas diferentes execuções do algoritmo; (2) os conjuntos $BestRules$ construídos possuem regras em comum mas não são idênticos. A Tabela 5 mostra o corpo e a cabeça (classe) das regras pertencentes aos conjuntos $BestRules$ construídos pelo \mathcal{RCO} , na sintaxe padrão Prolog, o número de vezes que a regra foi selecionada utilizando os seis critérios (#Sel.), e a opinião do especialista (Esp.) em relação á regra (**N** se não faz sentido, **OK** se a regra faz sentido e **?** se a regra não foi compreendida pelo especialista). Quando a regra foi rotulada com **N** na opinião do especialista, a tabela também

⁴Os classificadores \mathbf{h}_1 , \mathbf{h}_2 e \mathbf{h}_3 são distintos. Em (Bernardini & Monard 2002), são mostradas todas as regras pertencentes a esses classificadores.

⁵ se é o *standard error*, dado pelo desvio padrão dos erros encontrados dividido por $\sqrt{10}$. O valor 10 é utilizado devido ao 10 em 10-fold cross validation.

Tabela 3: Resultados Obtidos Utilizando o $\mathcal{CN}2$

Conj. de Dados → Classificador	EA	EC ($e \pm dp$ (%))	Erro ($e \pm se$ (%))	# Regras Induzidas
$S \rightarrow \mathbf{h}_{\mathcal{CN}2}$	12.90%	1 – 14.96±3.59 2 – 82.58±5.65 3 – 44.93±5.74	40.83±2.83	33
$S_1 \rightarrow \mathbf{h}_1$	3.30%	1 – 33.33±10.54 2 – 80.00±13.33 3 – 20.67±7.11	41.66±5.13	12
$S_2 \rightarrow \mathbf{h}_2$	1.70%	1 – 8.50±4.60 2 – 60.00±16.33 3 – 40.83±12.94	36.66±7.78	13
$S_3 \rightarrow \mathbf{h}_3$	0.00%	1 – 20.83±7.38 2 – 80.00±13.33 3 – 18.33±8.03	41.68±5.69	16

Tabela 4: Resultados Obtidos Utilizando o \mathcal{RCO}

Conjunto de Regras	# Regras Selecionadas	# Exemplos Restantes
BR_{accR}	12	26 – 43.33%
BR_{covR}	7	21 – 35.00%
BR_{satR}	14	21 – 35.00%
BR_{sensR}	6	24 – 40.00%
BR_{specR}	5	49 – 81.66%
BR_{supR}	6	25 – 41.67%

Tabela 5: Regras Seleccionadas pelo \mathcal{RCO}

B	H	# Sel.	Esp.	Razão
greater(idade,36.5), greater(class_B,29.5)	1	3	N	1
greater(tempo_abs,1.5), less(class_C,65), less(vitalidade,53.5)	1	3	N	3
greater(volume,4.1), greater(pH,8.25)	3	3	N	4
less(processamento,32.5), less(volume,4.2), less(concentracao_total,27.95)	1	3	OK	-
greater(pH,7.75), less(pH,8.25), less(concentracao_total,72.7), less(motilidade,12.5)	1	3	OK	-
greater(class_C,34.5), less(kruger,1.5)	1	2	OK	-
greater(idade,29), less(tempo_abs,6.5), less(hP,58.5)	1	2	?	-
less(idade,48.5), less(concentracao,33.65), greater(class_B,7), less(hP,75.5)	1	5	N	1 e 5
greater(vitalidade,70.5), greater(kruger,11.5), greater(hP,60)	3	2	OK	-
greater(concentracao_total,221.25), less(class_B,28.5), less(class_C,70.5)	3	5	OK	-
greater(concentracao,24), greater(motilidade,28.5)	3	2	OK	-
greater(volume,1.35), less(num_leu,1.28), less(hP,69)	1	3	N	-
greater(motilidade,19.5), less(num_leu,0.68)	3	1	N	7
less(pH,8.25), greater(concentracao,63.5), greater(motilidade,0.5)	3	4	OK	OBS. 6
less(class_A,11.5)	2	1	N	1
greater(class_B,25.5), greater(kruger,3.5)	2	1	OK	
greater(idade,31.5), less(processamento,35), greater(concentracao,20.35), less(class_A,21.5)	3	1	N	1

Tabela 6: Motivos do Especialista para Considerar que uma Regra Não Faz Sentido

Motivos do Especialista
1: Quanto maior o valor dos atributos class_A ou class_B, maior a chance da regra predizer as classes 2 ou 3
2: Quanto maior o valor dos atributos class_C ou class_D, menor a chance da regra predizer a classe 3
3: Quanto menor o valor dos atributos class_C ou class_D, menor a chance da regra predizer a classe 1
4: Especialista não utilizaria os atributos da regra
5: Valor de corte muito alto para o atributo hP
6: Valor de corte baixo para o atributo motilidade
7: Quanto menor o valor do atributo num_leu, menor a chance da regra predizer a classe 3

contém o índice do motivo que o levou a dizer que a regra não faz sentido (Razão). Os motivos do especialista e seus respectivos índices constam na Tabela 6.

Observa-se, na Tabela 5, que aproximadamente 50% das regras segue o raciocínio que o especialista utilizaria para analisar um caso. Os outros 50% não fazem sentido ora porque o *Corpo* da classe possui condições falsas, ora porque utiliza atributos que o especialista não acha suficientes para chegar a uma conclusão.

5 Conclusão

Neste trabalho é proposto o algoritmo \mathcal{RCO} , que tem como objetivo seleccionar o melhor conjunto de regras de conhecimento induzidas por algoritmos de aprendizado simbólico, segundo uma medida de avaliação de regras. O \mathcal{RCO} faz parte de um dos módulos do $\mathcal{R}_{ule}\mathcal{S}_{ystem}$, o qual é um sistema mais abrangente para auxiliar o usuário na análise de regras de conhecimento. Os resultados obtidos em um estudo de caso do mundo real utilizando o \mathcal{RCO} foram considerados interessantes tendo em vista, entre outros, que o número de regras a serem analisadas pelo especialista é menor que o número de regras iniciais. Além disso, o especialista considerou que bastantes regras apresentaram características que justificam a continuação do trabalho no domínio de processamento de sêmen. Atualmente, o especialista está colecionando mais casos com o objetivo de incrementar o conjunto de exemplos de treinamento, a fim de continuar a investigação inicial aqui apresentada.

Referências

- Baranauskas, J. A. & M. C. Monard (2000, Fev). Reviewing some machine learning concepts and methods. Technical Report 102, ICMC-USP. ftp://ftp.icmc.sc.usp.br/pub/BIBLIOTECA/rel_tec/Rt_102.ps.zip.
- Bernardini, F. C. (2002). Combinação de classificadores simbólicos para melhorar o poder preditivo e descritivo de *ensembles*. Dissertação de Mestrado, ICMC-USP.
- Bernardini, F. C. & M. C. Monard (2002). Um estudo de caso utilizando o módulo de explicação e combinação do $\mathcal{R}_{ule}S_{ystem}$. Technical Report 160, ICMC-USP. ftp://ftp.icmc.sc.usp.br/pub/BIBLIOTECA/rel_tec/Rt_160.ps.zip.
- Clark, P. & T. Niblett (1989). The $CN2$ induction algorithm. *Machine Learning* 3(4), 261–283.
- Freitas, A. A. (1998a). A multi-criteria approach for the evaluation of rule interestingness. In *Proceedings of the International Conference on Data Mining*, Rio de Janeiro, RJ, pp. 7–20.
- Freitas, A. A. (1998b). On objective measures of rule surprisingness. In *Proceedings of the Second European Symp, Lecture Notes in Artificial Intelligence*, Volume 1510, pp. 1–9.
- Gomes, A. K. (2002). Medidas de avaliação de regras. Dissertação de Mestrado, ICMC-USP.
- Hilderman, R. J. & H. J. Hamilton (1999). Knowledge discovery and interestingness measures: A survey. Technical Report s4s, University of Regina.
- Lavrac, N., P. Flach, & B. Zupan (1999). Rule evaluation measures: a unifying view. In *Proceedings of the Ninth International Workshop on Inductive Logic Programming, Lecture Notes in Artificial Intelligence*, Volume 1634, pp. 74–185.
- Lee, H. D. (2000). Seleção e construção de features relevantes para o aprendizado de máquina. Dissertação de Mestrado, ICMC-USP.
- Prati, R. C., J. A. Baranauskas, & M. C. Monard (2001). Extração de informações padronizadas para a avaliação de regras induzidas por algoritmos de aprendizado de máquina simbólico. Technical Report 145, ICMC-USP. ftp://ftp.icmc.sc.usp.br/pub/BIBLIOTECA/rel_tec/RT_145.ps.zip.