ISISTAN-UNCPBA

Curso: Aprendizado Simbólico de Máquina
Responsável: Profa. Maria Carolina Monard ICMC-USP
mcmonard@icmc.sc.usp.br
Datas: 9 a 20 de abril de 2001

LISTA 2

1 Classificadores Simbólicos Induzidos com o Conjunto de Dados de Treinamento da Lista 1

Os seguintes três classificadores, utilizando escolhas *pessoais* para escolher o atributo relevante em cada passo da construção do conceito induzido utilizando árvores de decisão, foram construí dos:

G1: Alfredo & Gustavo Tabela 1
G2: Flavia, Lucio, Mauricio & Pablo Tabela 2
G3: Curti, Valaco & Valdata Tabela 3

R01: IF humidity $\leq$ 78 AND temperature $\leq$ 20
THEN CLASS = dont_go

R02: IF humidity $\leq$ 78 AND temperature $>$ 20
THEN CLASS = go

R03: IF humidity $>$ 78 AND temperature $\leq$ 23 AND windy=yes
THEN CLASS = dont_go

R04: IF humidity $>$ 78 AND temperature $\leq$ 23 AND windy=no
THEN CLASS = go

R05: IF humidity $>$ 78 AND temperature $>$ 23 AND outlook=rain
THEN CLASS = go

R06: IF humidity $>$ 78 AND temperature $>$ 23 AND outlook=sunny
THEN CLASS = dont_go

R07: IF humidity $>$ 78 AND temperature $>$ 23 AND outlook=overcast
THEN CLASS = go

DEFAULT CLASS = go

Tabela 1: Regras Induzidas *voyage* G1

R01: IF outlook=rain AND windy=yes
THEN CLASS = dont_go

R02: IF outlook=rain AND windy=no
THEN CLASS = go

R03: IF outlook=sunny AND humidity ≥ 75
THEN CLASS = dont_go

R04: IF outlook=sunny AND humidity < 75
THEN CLASS = go

R05: IF outlook=overcast AND temperature ≥ 21
THEN CLASS = go

R06: IF outlook=overcast AND temperature < 21
THEN CLASS = dont_go

DEFAULT CLASS = go

Tabela 2: Regras Induzidas *voyage* G2

A Tabela 4 mostra as regras correspondentes a árvore de decisão induzida pelo algoritmo *C4.5* de Ross Quinlan

Lembre que as regras anteriores são regras de decisão. Calcule o erro total de cada classificador — Tabela 5 — usando

1. o proprio conjunto de treinamento (erro aparente)
2. o conjunto de teste

os quais encontram-se na Lista 1.

A Tabela 6 mostra as regras induzidas pelo algoritmo *C4.5rules* de Ross Quinlan. Lembre que essas regras são *unordered rules* dentro da mesma classe e *ordered* entre classes diferentes. Similar ao anterior, calcule o erro desse classificador.

Finalmente, a Tabela 7 mostra as regras induzidas pelo algoritmo *CN2* de Clark & Niblett. Lembre que essas regras são *unordered rules*. Primeiramente calcule, utilizando somente o conjunto de treinamento, o número de exemplos cobertos, por cada regra, da classe `go` e `dont_go`, *i.e.* [`go`,`dont_go`]. Logo após calcule o erro desse classificador.

Compare os resultados obtidos para cada classificador. Para esse conjunto de dados específico e somente levando em conta as experiências realizadas, qual classificador voce prefere? Justifique a sua resposta.

R01: IF windy=yes AND outlook=rain
THEN CLASS = go

R02: IF windy=yes AND outlook=sunny AND temperature ≤ 23
THEN CLASS = dont_go

R03: IF windy=yes AND outlook=sunny AND temperature > 23
THEN CLASS = go

R04: IF windy=yes AND outlook=overcast AND temperature ≤ 21
THEN CLASS = dont_go

R05: IF windy=yes AND outlook=overcast AND temperature > 21
THEN CLASS = go

R06: IF windy=no AND outlook=rain
THEN CLASS = go

R07: IF windy=no AND outlook=overcast
THEN CLASS = go

R08: IF windy=no AND outlook=sunny AND humidity < 79
THEN CLASS = go

R09: IF windy=no AND outlook=sunny AND humidity ≥ 79
THEN CLASS = dont_go

DEFAULT CLASS = go

Tabela 3: Regras Induzidas *voyage* G3

R1: IF outlook = overcast
THEN CLASS = go

R2: IF outlook = sunny AND humidity ≤ 78
THEN CLASS = go

R3: IF outlook = sunny AND humidity > 78
THEN CLASS = dont_go

R4: IF outlook = rain AND windy = yes
THEN CLASS = dont_go

R5: IF outlook = rain AND windy = no
THEN CLASS = go

DEFAULT CLASS = go

Tabela 4: Regras Induzidas *voyage* pelo Algoritmo *C4.5*

Rótulo da classe	predito como go	predito como dont_go	Taxa de Erro da Classe	Taxa de erro total
	Verd. positivos	Falsos negativos		
verd. go	T_P	F_N	$\frac{F_N}{T_P+F_N}$	$\frac{F_P+F_N}{n}$
	Falsos positivos	Verd. negativos		
verd. dont_go	F_P	T_N	$\frac{F_P}{F_P+T_N}$	

Tabela 5: Matriz de Confusão com $n = (T_P + F_N + F_P + T_N)$

R1: IF outlook = rain AND windy=no
THEN CLASS = go

R2: IF outlook = sunny AND humidity > 78
THEN CLASS = dont_go

R3: IF outlook = rain AND windy=yes
THEN CLASS = dont_go

DEFAULT CLASS = go

Tabela 6: Regras Induzidas *voyage* pelo Algoritmo *C4.5rules*

R1: IF humidity < 83.00 AND windy = no
THEN voyage = go [ ]

R2: IF outlook = overcast AND temperature > 19.50
THEN voyage = go [ ]

R3: IF outlook = sunny AND humidity < 76.00
THEN voyage = go [ ]

R4: IF outlook = rain AND humidity > 87.50
THEN voyage = go [ ]

R5: IF outlook = sunny AND humidity > 83.00
THEN voyage = dont_go [ ]

R6: IF temperature < 24.00 AND humidity < 80.50 AND windy = yes
THEN voyage = dont_go [ ]

DEFAULT CLASS = go [ ]

Tabela 7: Regras Induzidas *voyage* pelo Algoritmo *CN2*