



UNIDADE II: APRENDIZAGEM SUPERVISIONADA

- Sumário:

- Introdução
- Aprendizagem de conceitos e regras
 - Algoritmo AQ



Objectivos

- Adquirir a noção de regras de decisão
- Adquirir uma noção acerca do processo de aprendizagem de regras mediante estratégias de cobertura sequencial
- Descrever o algoritmo AQ



Aprendizagem de conceitos

- Consiste em encontrar a representação de um conceito, inicialmente desconhecido, a partir de um conjunto de exemplos e contra-exemplos do mesmo
- Pressupõe que os exemplos estão previamente etiquetados (classificados como exemplos positivos e negativos) → ***Aprendizagem supervisionada***
- Geralmente o conceito é expressado mediante um predicado ou propriedade que serve para caracterizar o espaço de objectos observado



Aprendizagem de conceitos

- Por exemplo o conceito “ser maçã” serve para discriminar se um objecto é ou não uma maçã
- A definição do conceito se descreve a partir das características que descrevem os objectos observados
 - Cor, tamanho, sabor, textura, densidade, tipo de alimento, país de origem...



Aprendizagem de regras de decisão

- Consiste em obter uma descrição geral de um conjunto de exemplos, permitindo a sua agrupação em dois grupos: instancias representantes do conceito buscado (exemplos) e que não são representantes do conceito buscado (contra-exemplos)
- A descrição deve ser mais geral que os exemplos observados, permitindo a classificação de novos exemplos



Aprendizagem de regras de decisão

- O conceito aprendido se representa através de um conjunto de regras com a forma:
 - $R_1 : \textit{antecedente}_1$ então c_1
 - $R_2 : \textit{antecedente}_2$ então c_1
 - ...
 - $R_k : \textit{antecedente}_k$ então c_1
 - $R_{k+1} : \textit{antecedente}_{k+1}$ então c_2
 - ...
 - $R_m : \textit{antecedente}_m$ então c_{n-1}
 - senão c_n



Aprendizagem de regras de decisão

- A tarefa de aprendizagem se pode definir como:
 - Dados:
 - Um conjunto de instâncias positivas (I^+) e/ou negativas (I^-) do conceito buscado
 - Determinar uma hipótese (h) que:
 - Descreva todas (ou a maioria) das instâncias positivas
 - Exclua todas (ou a maioria) das instâncias negativas
 - Tenha expectativas de classificar correctamente instâncias ainda não analisadas



Aprendizagem de regras de decisão

- Enfoque seguido consiste em especificar a tarefa como uma ***busca de uma hipótese*** desconhecida num espaço de hipóteses a partir dos exemplos disponíveis
- A estratégia seguida consiste em considerar uma classe de cada vez e buscar uma forma de cobrir todas as instâncias pertencentes à mesma e simultaneamente excluir todas as demais instâncias (*sequential covering*)



Algoritmo AQ

- Constitui uma família de métodos simbólicos derivados do método ***estrela*** introduzido por Ryszard Michalski (1937 – 2007) nos anos 70
- O objectivo do algoritmo consiste em obter um conjunto de regras de classificação que descrevam todos os exemplos positivos de um conjunto de dados e não descrevam nenhum negativo



Algoritmo AQ

- Entradas
 - Conjunto de dados de treino
 - Função lexicográfica de avaliação (LEF)
- Saídas
 - Conjunto de regras gerais que cobrem todas as instâncias positivas e nenhuma negativa



Função Lexicográfica de Avaliação (LEF)

- A função LEF estabelece uma lista de critérios de preferência de regras
- Toma como entradas um conjunto de regras candidatas à descrição de um exemplo (uma **estrela**) e selecciona a mais apropriada, tendo em conta uma série de critérios de preferência



Função Lexicográfica de Avaliação (LEF)

- Os critérios podem ser entre outros
 - Cobertura: quantidade de exemplos positivos cobertos pela regra
 - Simplicidade: número de atributos que aparecem nas condições da regra
 - Generalidade: estimada como o número exemplos observados que a regra descreve, dividido pelo número de possíveis exemplos...



Saídas

- Para a representação das saídas utiliza um tipo de cálculo proposicional (VL1)
- Principais conceitos relacionados:
 - ***Selector***: permite realizar perguntas sobre os valores dos atributos de entrada. Sintaxe (Atributo Operador Valor)
 - ***Complexo (regra)***: conjunção de selectores
 - ***Recobrimento (cover)***: disjunção de complexos



Descrição

- A tarefa se estrutura em duas buscas
 - Na mais externa se busca um conjunto de regras que classifique correctamente a todos os exemplos
 - Na interna, para cada exemplo positivo se busca um conjunto de regras que descrevam ao mesmo e não descreva a nenhum exemplo negativo



Descrição

- Processo levado a cabo de forma iterativa
- Em cada passo
 - Escolhe-se do conjunto de entrada um exemplo positivo, denominado ***semente*** (*seed*)
 - Gera-se mediante a busca interna um conjunto de complexos que o descrevem e não descrevem a nenhum exemplo negativo (conjunto designado ***estrela***)
 - Escolhe-se qual desses complexos incluir no recobrimento formado até ao momento, através da função LEF
 - Remove-se do conjunto todos os exemplos positivos cobertos pelo complexo seleccionado



Algoritmo

- Função *Estrela*
- Entradas
 - P : conjunto de exemplos positivos
 - N : conjunto de exemplos negativos
 - LEF : função lexicográfica
- Saída
 - R : conjunto de regras
- $R := \emptyset$
- Enquanto $P \neq \emptyset$
 - $semente := escolher_exemplo(P)$
 - $estrela := determina_estrela(semente, N)$
 - $complexo := escolher_complexo(estrela, P, N, LEF)$
 - $R := R \cup \{complexo\}$
 - $P := P - exemplos_cobertos_por(P, complexo)$
- Devolver R

Algoritmo

- Função determina_estrela
- Entradas
 - p : exemplo positivo
 - N : conjunto de exemplos negativos
- Saída
 - E : estrela ou conjunto de regras
- $E := \phi$
- $L := (\Pi)$
- S := gerar_todos_os_selectores(p)
- Enquanto $L \neq \phi$
 - $L' := \{x \wedge y \mid x \in L, y \in S\} - \{E \cup L\}$
 - Para cada complexo $C_i \in L'$
 - Se válido(C_i , N) Então
 - $E := \{C_i\} \cup E$
 -
 - $L := L'$
- Devolver estrela E

Exemplo

#	Ante nas	Cau das	Núc leos	Corpo	Classe
1	1	0	2	Riscas	Normal
2	1	0	1	Branco	Cancerígena
3	1	2	0	Riscas	Normal
4	0	2	1	Riscas	Normal
5	1	1	1	Riscas	Cancerígena
6	2	2	1	Riscas	Cancerígena
7	0	2	2	Riscas	Normal

- $S = \{(A_1=1), (A_2=0), (A_3=2), (A_4=Riscas)\}$

...

~~$C_{11} : A_1 = 1 \Rightarrow N$~~

~~$C_{12} : A_2 = 0 \Rightarrow N$~~

$C_{13} : A_3 = 2 \Rightarrow N$

~~$C_{14} : A_4 = R \Rightarrow N$~~

...

$C_{111} : A_1 = 1 \wedge A_2 = 0 \Rightarrow N$

$C_{112} : A_1 = 1 \wedge A_3 = 2 \Rightarrow N$

$C_{113} : A_1 = 1 \wedge A_4 = R \Rightarrow N$

...

Exemplo

#	Ante nas	Cau das	Núc leos	Corpo	Classe
1	1	0	2	Riscas	Normal
2	1	0	1	Branco	Cancerígena
3	1	2	0	Riscas	Normal
4	0	2	1	Riscas	Normal
5	1	1	1	Riscas	Cancerígena
6	2	2	1	Riscas	Cancerígena
7	0	2	2	Riscas	Normal

...

$$C_{13} : A_3 = 2 \Rightarrow N$$

~~$$C_{1111} : A_1 = 1 \wedge A_2 = 0 \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{1112} : A_1 = 1 \wedge A_2 = 0 \wedge A_4 = R \Rightarrow N$$~~

~~$$C_{112} : A_1 = 1 \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{1132} : A_1 = 1 \wedge A_4 = R \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{122} : A_2 = 0 \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{123} : A_2 = 0 \wedge A_4 = R \Rightarrow N$$~~

~~$$C_{143} : A_4 = R \wedge A_3 = 2 \Rightarrow N$$~~

LEF = {(cobertura, 1), (# premissas, 2)}

$$R = \{C_{13}\}$$



Exemplo

- LEF = {(cobertura, 1), (# premissas, 3)}

$$C_{13} : A_3 = 2 \Rightarrow N$$

~~$$C_{1111} : A_1 = 1 \wedge A_2 = 0 \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{1112} : A_1 = 1 \wedge A_2 = 0 \wedge A_4 = R \Rightarrow N$$~~

~~$$C_{112} : A_1 = 1 \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{1132} : A_1 = 1 \wedge A_4 = R \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{122} : A_2 = 0 \wedge A_3 = 2 \Rightarrow N$$~~

~~$$C_{123} : A_2 = 0 \wedge A_4 = R \Rightarrow N$$~~

~~$$C_{143} : A_4 = R \wedge A_3 = 2 \Rightarrow N$$~~

$$R = \{C_{13}\}$$

#	Ante nas	Cau das	Núc leos	Corpo	Classe
2	1	0	1	Branco	Cancerígena
3	1	2	0	Riscas	Normal
4	0	2	1	Riscas	Normal
5	1	1	1	Riscas	Cancerígena
6	2	2	1	Riscas	Cancerígena



Vantagens

- As descrições geradas são simbólicas, sendo úteis quando se deseja que os utilizadores conheçam uma descrição de alto nível dos exemplos
- O utilizador pode proporcionar como entrada que tipo de descrição prefere, através da função LEF
- É possível definir diferentes estratégias de busca da estrela, de forma que se pode tornar mais eficiente o processo de busca



Desvantagens

- Na sua versão original, não trata com exemplos com atributos contínuos ou com dados com ruído



Variantes

- Várias variantes:
 - Umas realizam tratamento de atributos com valores contínuos
 - Outras diferem na forma em que se realiza a busca interna
 - Outras geram conjuntos de regras ordenadas (listas de decisão)...
- CN2, AQ11, AQ15, ...



Tarefa

- Seguir exemplo e verificar qual deve ser o recobrimento final



Bibliografia

- Mitchell, pg. 20 – 25
- Borrajo Millán
- Sierra Araujo
- Leitura da semana
 - Michalski, The AQ Family of Learning Programs, Ler abstract, Epígrafes 1 e 2. Epígrafe 3 ver de maneira geral diferentes variantes