



K-NN

1. Considere o seguinte conjunto de treino:

ID	a1	a2	a3	classe
1	1	1	0	X
2	0	1	1	Z
3	0	0	1	X
4	1	1	1	X
5	0	1	0	X
6	1	0	1	Z

Classifique o exemplo $x = \{a1=1, a2=0, a3=1\}$ usando:

- a) classificador 1-NN (distância Euclideana);
- b) classificador 5-NN (distância Euclideana).



Naive Bayes

1. Considere o seguinte conjunto de treino:

ID	a1	a2	a3	classe
1	1	1	0	X
2	0	1	1	Z
3	0	0	1	X
4	1	1	1	X
5	0	1	0	X
6	1	0	1	Z

Classifique o exemplo $x = \{a1=0, a2=1, a3=1\}$ usando um *Naive Bayes*.

2. Considere os dados do ficheiro "golf.csv".

Outlook	Temperature	Humidity	Wind	Play
sunny	85	85	FALSE	no
sunny	80	90	TRUE	no
overcast	83	78	FALSE	yes
rain	70	96	FALSE	yes
rain	68	80	FALSE	yes
rain	65	70	TRUE	no
overcast	64	65	TRUE	yes
sunny	72	95	FALSE	no
sunny	69	70	FALSE	yes
rain	75	80	FALSE	yes
sunny	75	70	TRUE	yes
overcast	72	90	TRUE	yes
overcast	81	75	FALSE	yes
rain	71	80	TRUE	no

Classifique o exemplo $x = \{\text{Outlook} = \text{"sunny"}, \text{Temperature} = 66, \text{Humidity} = 90, \text{Wind} = \text{TRUE}\}$ usando *Naive Bayes*.

Árvores de Decisão e Regras de Decisão

1. Fazer o diagrama em árvore da disjunção (i.e. OR).

A	B	V
FALSE	FALSE	FALSE
FALSE	TRUE	TRUE
TRUE	FALSE	TRUE
TRUE	TRUE	TRUE

2. Considere o conjunto de dados apresentado na tabela seguinte em que **DMC** é a duração média de uma chamada, **FUM** é a faturação do último mês, e **CT** é uma variável que indica se o contrato terminou. A partir destes três atributos pretende-se determinar o valor do atributo **Ab** (abandonou?)

DMC	FUM	CT	Ab
alta	alta	sim	sim
alta	baixa	não	não
baixa	alta	não	não
baixa	baixa	sim	sim
baixa	alta	sim	não

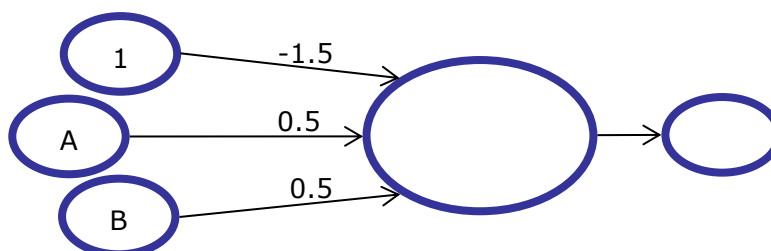
Extraia a Regra de Decisão utilizando o algoritmo oneR.

3. Considere a árvore criada anteriormente. Expand a folha "impura" utilizando o oneR, e considerando apenas as variáveis DMC e FUM.
4. Da árvore criada, extraia um conjunto de regras com o algoritmo *top-down* (STOP: cobertura = 80%, qualidade = 100%).
5. Usando o algoritmo *bottom-up*, e a partir da regra "**IF** CT = sim & FUM = alto **THEN** Ab = sim", extraia uma regra com cobertura $\geq 60\%$ e qualidade $\geq 60\%$.

Redes Neurais

1. Considere o seguinte conjunto de dados e o *perceptron*

A	B	\wedge
0	0	0
0	1	0
1	0	0
1	1	1



1.1. Utilizando a taxa de aprendizagem de $\eta = 0.25$ e a função de ativação,

$$f(x_1, x_2, \dots, x_n) = \begin{cases} 1, & \text{sse } \sum x_i w_i > 0 \\ 0, & \text{c.c.} \end{cases},$$

descreva uma iteração do algoritmo de correção de erro do *perceptron* para o conjunto de testes da tabela (função AND).

1.2. O algoritmo convergiu?



Support Vector Machine

1. Considere os seguintes pontos:

x	y	classe	x	y	classe
2	5	•	1	1	■
5	6	•	2	3	■
3	7	•	4	2	■
6	8	•	5	4	■

1.1. Qual o hiperplano paralelo ao eixo do x com a maior margem?

1.2. Este será o melhor hiperplano?

Algoritmos Genéticos

1. Crie 10 instâncias para o “problema no *Knapsack*”. Considere uma taxa de seleção de 50% (as melhores 5 instâncias contribuem com 2 cópias cada). Como *crossover*, escolha 3 pares diferentes das instâncias selecionadas e troque os últimos 3 genes. Finalmente, insira uma mutação aleatória em cada instância. Repita o processo 2 vezes e indique a solução final.

ITEM	SURVIVAL POINTS	WEIGHT
pocketknife	10.00	1.00
beans	20.00	5.00
potatoes	15.00	10.00
unions	2.00	1.00
sleeping bag	30.00	7.00
rope	10.00	5.00
compass	30.00	1.00