

## Problemas práticos de Datamining II – Modelos preditivos

### I

Imagine que havia o risco de termos que empenhar uma fragata num cenário operacional onde havia uma séria ameaça aérea, mas onde havia também muitas aeronaves amigas. A distinção entre os dois é por isso fundamental. Durante exercícios navais foi possível observar em acção as aeronaves amigas e, graças a um esforço de “shadowing” das forças adversárias, foi possível observar, ainda que muito poucas vezes, aeronaves adversárias em acção. No entanto, durante a operação em causa, estima-se que a probabilidade de encontrar aeronaves adversárias seja substancialmente mais alta.

A experiência observada durante os exercícios foi a seguinte (1=sim, 0=não):

Característica	nº da observação	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A - Voava abaixo dos 200m		1	1	0	0	1	0	0	1	0	1	1	0	0	1	0	1	0	0	0	0	0	1	0	1
B - Voava abaixo dos 800m		1	1	1	1	1	0	1	1	0	1	1	1	1	1	0	1	1	1	0	0	1	1	0	1
C - Emitia na banda K		0	0	0	1	1	1	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	1	0	1
D - Voava acima de mach 1		0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	1	1	1	0	0	0	0
E - Tinha o radar em busca cónica		1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0
T - Era adversário		1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

- 1) Se, durante os exercícios navais, uma das fragatas detectar um contacto com o radar em busca cónica e a voar abaixo dos 800m de altitude, deve classificá-lo como amigo ou inimigo ?
- 2) Se se assumir que nas operações reais é igualmente provável encontrar aeronaves amigas e inimigas, e se durante essas operações uma das fragatas detectar um contacto com o radar em busca cónica e a voar abaixo dos 800m de altitude, deve classificá-lo como amigo ou inimigo ?

### II

Usando a base de dados de seguros (ver problema prático),

Aparece um cliente que tem claramente mais de 35 anos, mas menos que 65, vem acompanhado da esposa e do filho, tem carta e seguro nesta companhia há muitos anos, mas apenas do automóvel contra terceiros, tem curso superior, não é estudante, tem casa própria e múltiplos telemóveis, e fuma. Deve tentar fazer-lhe um seguro ?

Induza uma árvore de decisão a partir dos dados disponíveis, mas usando apenas os campos m35, m65, cas, tc5, sm,,est,tmt, e fum. O cliente acima é um bom ou mau cliente ?