

**Práctico 3**

Aprendizaje bayesiano. Aprendizaje por casos.

**Ejercicio 1**

Suponga que quiere implementar un detector de correo no deseado y cuenta con el siguiente conjunto de entrenamiento:

#	Agendado	Idioma	Para	'Venta'	Deseado
1	Sí	Inglés	Sí	Sí	Sí
2	No	Otro	No	No	Sí
3	No	Español	No	Sí	Sí
4	Sí	Otro	Sí	No	No
5	No	Español	Sí	Sí	No

Considere el ejemplo:

#	Agendado	Idioma	Para	'Venta'	Deseado
6	Sí	Inglés	No	Sí	??

- ¿Cómo clasificaría un clasificador bayesiano sencillo al correo #6? ¿Con qué probabilidad?
- ¿Qué modificación debería hacer en la parte anterior, si le informan que el atributo «Idioma» puede tomar el valor «francés»?
- ¿Cómo implementaría un 2-NN? Ejemplifique su solución clasificando a #6.

**Ejercicio 2**

Suponga que el doctor del ejemplo visto en el teórico decide realizar un segundo test a su paciente y vuelve a dar positivo. Asumiendo que ambos tests son independientes, ¿cuáles son ahora las probabilidades de que el paciente tenga o no tenga cáncer?

**Ejercicio 3**

Suponga que cuenta con un conjunto de entrenamiento  $D$  sin ruido, y se considera un espacio  $H$  en donde las hipótesis, cuanto más generales, mayor probabilidad tienen de ser el concepto objetivo. Indique si las siguientes afirmaciones son verdaderas o falsas. Justifique.

- Find-S da una hipótesis MAP.
- Find-S da una hipótesis ML.
- Halving es un clasificador bayesiano óptimo.
- Si  $VS_{H,D}$  es vacío, se debe cambiar el espacio considerado.

**Ejercicio 4**

Considere el algoritmo *Find-G*, que da como resultado alguna de las hipótesis más genéricas del espacio de versiones. Dé distribuciones para  $P(h)$  y  $P(D|h)$  bajo las cuales:

- Find-G da una hipótesis MAP.
- Find-G puede no dar hipótesis MAP.
- Find-G seguramente da una hipótesis ML pero no MAP.

**Ejercicio 5**

a) Usted cuenta con un conjunto de entrenamiento con ruido. ¿Qué ventajas y desventajas encuentra al aplicar los siguientes algoritmos? ¿Cuál de éstos elegiría?

- Candidate-Elimination
- ID3
- Clasificador bayesiano óptimo
- K-NN

b) ¿Cambiaría su respuesta si el conjunto no tiene ruido pero sólo tiene instancias positivas?

**Ejercicio 6**

Se desea aplicar el principio MDL a un espacio de conjunciones de hasta  $n$  atributos booleanos; por ejemplo: *Soleado*  $\wedge$  *SinCambios*. Se tiene un conjunto de  $m$  ejemplos.

Cada hipótesis se transmite listando sus atributos; cada atributo se codifica utilizando  $\log(n)$  bits. Dada una hipótesis  $h$ , la codificación de un ejemplo tiene largo cero si  $h$  lo clasifica correctamente, y largo  $\log(m)$  en caso contrario (para indicar cuál ejemplo es errado).

- a) Dé la expresión a minimizar según el principio MDL.
- b) ¿Es posible construir un conjunto de entrenamiento que tenga una hipótesis consistente pero haga que MDL elija otra menos consistente? Justifique.
- c) Plantee distribuciones para  $P(h)$  y  $P(D|h)$  bajo las cuales el algoritmo MDL da como resultado una hipótesis MAP.

**Ejercicio 7 [\*]**

Considere el ejercicio 6 del práctico anterior y resuélvalo utilizando los siguientes algoritmos en lugar del algoritmo ID3:

1. Clasificador bayesiano sencillo
2. K-NN, con  $k$  igual 1, 3 y 5.

Presente una evaluación detallada de los resultados. Compare los resultados de los tres algoritmos utilizados para resolver el problema.