

Apellidos:

Nombre:

Ejercicio 1 [2 puntos] **Cuestiones**

Contestar a las siguientes cuestiones *utilizando para ello* el espacio en blanco que se deja a continuación. No serán tenidas en cuenta las respuestas que se escriban fuera de ese espacio.

- Considérese la siguiente base de conocimiento BC :

R1: $g(x,y) \rightarrow p(x,y)$.R2: $g(x,z), p(z,y) \rightarrow p(x,y)$.H1: $g(a,c)$.H2: $g(b,c)$.H3: $g(c,e)$.H4: $g(d,e)$.

donde g y p son símbolos de función binarios y a, b, c, d y e son constantes. Dibujar el árbol SLD generado por SLD-RESOLUCION(BC, P) e indicar la(s) respuesta(s) encontrada(s), siendo P el objetivo $p(x, e)$.

- Consideremos la siguiente base de conocimiento en CLIPS:

```
(defacts hechos
  (G a c)
  (G b c)
  (G c e)
  (G d e))

(defrule regla-1
  (G ?x ?y)
  =>
  (assert (P ?x ?y)))

(defrule regla-2
  (G ?x ?z)
  (P ?z ?y)
  =>
  (assert (P ?x ?y)))
```

Construir una tabla de seguimiento de la ejecución del programa anterior ¿Qué se obtiene como resultado de la ejecución? ¿Cómo podríamos modificar el programa para que diera resultados similares al del apartado anterior?

Apellidos:

Nombre:

Ejercicio 1 (cont.)

- ¿Qué entendemos por inferencia probabilística? ¿Qué es una red bayesiana? ¿Qué relaciones de independencia se asumen en una red bayesiana? Demostrar que la información contenida en una red bayesiana, junto con las mencionadas relaciones de independencia asumidas, nos permiten calcular cualquier inferencia probabilística.

- Describir en pseudocódigo el algoritmo k -medias ¿Para qué se usa? ¿Por qué puede verse como un algoritmo de búsqueda local? ¿Qué se busca y qué se trata de optimizar?

- Describir el problema general de clasificación de documentos, algunas de sus aplicaciones y un algoritmo (en pseudocódigo) que use *Naive Bayes* usado para clasificación de documentos.

Apellidos:
 Nombre:

Ejercicio 2 [1.5 puntos] **Aprendizaje inductivo**

Se han encontrado una gran cantidad de obras de arte realizadas por dos artistas **A** y **B**, pero sólo para un pequeño número de obras se ha podido asegurar cuál de los dos es el autor.

La siguiente tabla muestra los datos de dichas obras, indicando el autor en función del TIPO de obra (grabado, óleo o acuarela), del LUGAR donde se encontró la obra (España, Portugal o Francia), de su ESTILO (clásico o moderno), y de si tienen MARCO o no.

Ejemplo	TIPO	LUGAR	ESTILO	MARCO	AUTOR
E_1	grabado	España	clásico	no	B
E_2	grabado	España	moderno	no	B
E_3	grabado	Portugal	moderno	no	B
E_4	grabado	Francia	clásico	si	B
E_5	grabado	Francia	moderno	no	B
E_6	grabado	Francia	moderno	si	B
E_7	óleo	España	clásico	si	A
E_8	óleo	España	clásico	no	A
E_9	óleo	Francia	moderno	no	A
E_{10}	óleo	Portugal	moderno	si	B
E_{11}	óleo	España	moderno	si	B
E_{12}	acuarela	Francia	clásico	no	B
E_{13}	acuarela	España	clásico	si	A
E_{14}	acuarela	Francia	moderno	no	B
E_{15}	acuarela	España	moderno	no	A
E_{16}	acuarela	Portugal	moderno	si	B

1. Aplicar (detallando cada uno de los pasos realizados) el **algoritmo ID3** para encontrar un árbol de decisión consistente con el conjunto de entrenamiento $\{E_1, \dots, E_{16}\}$ que permita decidir si una obra de arte fue realizada por **A** o por **B**.
2. Consideremos la siguiente tabla de ejemplos como conjunto de prueba
 - (a) calcular el rendimiento del árbol de decisión obtenido en el apartado anterior.
 - (b) aplicar (detallando los pasos realizados) un proceso de poda sobre dicho árbol.

Ejemplo	TIPO	LUGAR	ESTILO	MARCO	AUTOR
E_{17}	grabado	España	moderno	si	A
E_{18}	óleo	Portugal	moderno	no	A
E_{19}	óleo	Francia	moderno	si	B
E_{20}	óleo	España	moderno	no	A
E_{21}	acuarela	España	clásico	no	A
E_{22}	acuarela	Francia	clásico	si	B
E_{23}	acuarela	España	moderno	si	A
E_{24}	acuarela	Portugal	clásico	si	B

Tabla de Entropías $Ent(X, Y)$

		Y									
		0	1	2	3	4	5	6	7	8	9
X	0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	1	0.000	1.000	0.918	0.811	0.722	0.650	0.592	0.544	0.503	0.469
	2	0.000	0.918	1.000	0.971	0.918	0.863	0.811	0.764	0.722	0.684
	3	0.000	0.811	0.971	1.000	0.985	0.954	0.918	0.881	0.845	0.811
	4	0.000	0.722	0.918	0.985	1.000	0.991	0.971	0.946	0.918	0.890
	5	0.000	0.650	0.863	0.954	0.991	1.000	0.994	0.980	0.961	0.940
	6	0.000	0.592	0.811	0.918	0.971	0.994	1.000	0.996	0.985	0.971
	7	0.000	0.544	0.764	0.881	0.946	0.980	0.996	1.000	0.997	0.989
	8	0.000	0.503	0.722	0.845	0.918	0.961	0.985	0.997	1.000	0.998
	9	0.000	0.469	0.684	0.811	0.890	0.940	0.971	0.989	0.998	1.000

Apellidos:
Nombre:

Ejercicio 3 [1.5 puntos] **Aprendizaje estadístico** Una fábrica tiene como suministradoras de tornillos a las empresas A , B , C y D . De cada 100 cajas de tornillos que llegan a la fábrica, 75 son suministradas por la empresa A , 10 por la empresa B , 10 por la empresa C y 5 por la empresa D .

Además, sabemos que algunos tornillos son defectuosos y que la proporción de tornillos defectuosos depende de la empresa suministradora.

- El 10% de los tornillos suministrados por la empresa A son defectuosos.
- El 20% de los tornillos suministrados por B son defectuosos.
- El 30% de los tornillos suministrados por C son defectuosos
- El 50% de los tornillos suministrados por D son defectuosos.

Abrimos una caja y de los 10 primeros tornillos que tomamos, 4 de ellos son defectuosos. Queremos calcular la probabilidad de que el siguiente tornillo que tomemos de la caja sea defectuoso por diferentes métodos.

- Usando la *hipótesis de máxima verosimilitud* [0.5 puntos].
- Usando la *hipótesis más probable a posteriori* [0.5 punto].
- Usando *aprendizaje bayesiano* [0.5 puntos].

Explicar detalladamente cada uno de los pasos realizados.

Nota: Para el cálculo de la verosimilitud de los datos se tendrá en cuenta que coger un tornillo es un evento idénticamente distribuido e independiente de los anteriores.

Apellidos:
Nombre:

Ejercicio 4 [1.5 puntos] Redes Neuronales

Supongamos que queremos diseñar un sistema automatizado que regule la velocidad de giro de una fresadora en función de ciertos parámetros obtenidos a partir de cuatro sensores. Cada sensor genera datos en el intervalo $(0, 1)$ y la velocidad de giro se puede disminuir o incrementar hasta en un 50% con respecto a la velocidad original. Se sabe que la relación entre los datos proporcionados por los sensores y la regulación de la velocidad de giro de la fresadora no es lineal.

Responder detalladamente a las siguientes preguntas:

1. ¿Cómo se diseñaría este sistema mediante una red neuronal? Dibujar la red neuronal y explicar cómo se relaciona cada uno de sus elementos con la descripción del problema. ¿Qué función de activación habría que considerar?
2. ¿Cómo se realizaría el entrenamiento de la red neuronal anterior? Escribir las fórmulas de actualización de pesos para dicha red si utilizásemos el algoritmo de retropropagación.
3. ¿Cómo utilizarías dicha red en una situación concreta? Indica cómo modificarías la velocidad de giro de la máquina a partir de la salida generada por la red neuronal.