

Hoja 7 (Tema 4. Aprendizaje – I)

1) Define un problema y resuélvelo con el enfoque de Winston (mundo de bloques). Para ello:

a) Analiza la situación, determinando:

- 1.- El problema: crear una descripción textual de la definición a obtener
- 2.- Elementos relevantes: objetos, relaciones o propiedades (hacer listas)
- 3.- Determinar la representación formal de la descripción textual (ej.: ver trans 21)
- 4.- Construir ejemplos “+” y “-” (50% cada) en el lenguaje formal distinguiendo claramente el tipo: 1.- Buenos, 2.- Con problemas de ruido 3.- Malos, 4.- Sesgados 5.- Poco representativos 6.- Con atributos irrelevantes

b) Obtener una definición estructural siguiendo el Algoritmo de Winston, indicando *en cada paso* qué operación se ha aplicado: Variabilizar, Jerarquizar, Debilitar, Fortalecer o Cuantificar para todos.

2) Se quiere aplicar el algoritmo del espacio de versiones (Mitchell) para determinar la causa de la reacción alérgica que se produce en un paciente, en función del restaurante en el que ha comido, de si ha sido desayuno o comida, del día de la semana y del precio de la comida. Los datos de los que se dispone son los siguientes:

<i>Restaurante A</i>	<i>desayuno</i>	<i>viernes</i>	<i>barato</i>	<i>reacción +</i>
<i>Restaurante B</i>	<i>comida</i>	<i>viernes</i>	<i>caro</i>	<i>reacción -</i>
<i>Restaurante A</i>	<i>comida</i>	<i>sábado</i>	<i>barato</i>	<i>reacción +</i>
<i>Restaurante C</i>	<i>desayuno</i>	<i>domingo</i>	<i>barato</i>	<i>reacción -</i>
<i>Restaurante A</i>	<i>desayuno</i>	<i>domingo</i>	<i>caro</i>	<i>reacción -</i>

3) Aplicar el algoritmo ID3 para obtener el árbol de decisión que permita clasificar a las personas según su riesgo de padecer quemaduras solares. Los datos de los que se dispone son los siguientes:

<i>Nombre</i>	<i>Pelo</i>	<i>Estatura</i>	<i>Peso</i>	<i>Loción</i>	<i>Quemaduras</i>
<i>Sara</i>	<i>rubio</i>	<i>media</i>	<i>bajo</i>	<i>no</i>	<i>sí</i>
<i>Diana</i>	<i>rubio</i>	<i>alta</i>	<i>medio</i>	<i>sí</i>	<i>no</i>
<i>Alejandro</i>	<i>castaño</i>	<i>baja</i>	<i>medio</i>	<i>sí</i>	<i>no</i>
<i>Ana</i>	<i>rubio</i>	<i>baja</i>	<i>medio</i>	<i>no</i>	<i>sí</i>
<i>Emilio</i>	<i>pelirrojo</i>	<i>media</i>	<i>alto</i>	<i>no</i>	<i>sí</i>
<i>Pedro</i>	<i>castaño</i>	<i>alta</i>	<i>alto</i>	<i>no</i>	<i>no</i>
<i>Juan</i>	<i>castaño</i>	<i>media</i>	<i>alto</i>	<i>no</i>	<i>no</i>
<i>Carolina</i>	<i>rubio</i>	<i>baja</i>	<i>bajo</i>	<i>sí</i>	<i>no</i>

4) Definir un problema sobre un dominio real sobre el que conozcas ejemplos reales para aplicar el algoritmo ID3:

a) Analiza la situación, haciendo lo siguiente:

- 1.- Define varias clases a las que pertenecen los ejemplos (cada ejemplo pertenece a una de ellas). Esto es en vez de solo “+” y “-”
- 2.- Define qué atributos (y sus valores posibles) son relevantes para representar esos ejemplos.
- 3.- Estudia qué situaciones se pueden dar utilizando esos atributos, para describir ejemplos que cubran la mayoría de dichas situaciones.

4.- Define los ejemplos, que sean representativos y cubran las situaciones posibles que se pueden dar.

b) Genera el árbol de decisión aplicando el algoritmo ID3.

5) Se quiere construir un sistema de reconocimiento de objetos que utilice los resultados de un programa de visión que extrae características de los objetos a partir de la imagen obtenida por una cámara. De cada objeto se extraen tres características: tamaño, forma y número de orificios. Para entrenar al sistema, un tutor humano ha etiquetado cada ejemplo con el nombre del objeto al que corresponde, obteniendo los siguientes casos de entrenamiento:

<i>tuerca (pequeño, compacta, 1)</i>	<i>lápiz (grande, alargada, 0)</i>
<i>tijeras (grande, alargada, 2)</i>	<i>tornillo (pequeño, alargada, 0)</i>
<i>lápiz (grande, alargada, 0)</i>	<i>llave (pequeño, alargada, 1)</i>
<i>tijeras (grande, otras, 2)</i>	<i>llave (grande, alargada, 1)</i>
<i>llave (pequeño, otras, 2)</i>	<i>tornillo (pequeño, compacta, 0)</i>

Aplicar el algoritmo ID3 para construir el primer nivel del árbol de decisión, dejando indicadas las operaciones pero estableciendo cuáles serían los nodos del primer nivel en función del valor numérico de los resultados. Describir cómo se construiría el resto del árbol y cuándo se considerará que un nodo es terminal.

¿Qué diferencias habría si se aplicara el espacio de versiones a este mismo problema?

Describe cómo lo harías y qué ventajas y/o inconvenientes tendría.

6) Definir un problema sobre un dominio real sobre el que conozcas ejemplos reales para aplicar el algoritmo de clustering. NO debes saber intuitivamente cómo se clasificarán los objetos.

a) Analiza la situación, haciendo lo siguiente:

- 1.- Define: a)- Los atributos, y sus valores posibles, relevantes para representar esos ejemplos. b)- La distancia entre los valores de los atributos; cuando no son cuantitativos usa una heurística. c)- Define el orden de importancia de los atributos en relación con la relevancia para representar el objeto, también es heurístico porque no se sabe a priori.
- 2.- Define la distancia entre objetos basada en las distancias de cada atributo y la importancia de cada atributo.
- 3.- Define, usando esos atributos el conjunto de ejemplos (al menos 10).
- 4.- Define la Matriz de Distancias entre dichos ejemplos, aplicando la distancia definida en 4.

b) Aplicar el algoritmo de clustering y pintar el árbol de clasificación.

c) Conclusiones: qué conocimiento nuevo te ha aportado ese árbol respecto a los ejemplos que tienes, qué puedes decir nuevo sobre ellos?

7) Se desea obtener una descripción estructural del concepto “copa” aplicando deducción EBL. Se cuenta con la descripción funcional del concepto COPA representado por la regla C que se incluye en las reglas de la teoría completa del dominio:

Regla C (concepto objetivo):

copa (x): líquido_en (x),
estable (x),
beber_en (x)

Regla Le:

levantable (w): ligero (w),
sujetable (w).

Regla Li:

Líquido_en (u): material (u, v),
no_poroso (v),

Regla Su1:

sujetable (t): pequeño (t),
con_asas (t)

Regla Be:

beber_en (y): levantable (y),
cóncavo (y)

Regla Su2:

sujetable (s): pequeño (s),
cilíndrico (s)

Regla Es:

estable (z): pie_plano (z)

Hechos

ligero (porcelana),
ligero (vidrio),
ligero (aluminio)
no_poroso (porcelana),
no_poroso (vidrio),
no_poroso (aluminio)

Se cuenta también con la descripción del siguiente ejemplo positivo:

copa(copa1)
material(copa1,vidrio),pie_plano(copa1)
pequeño(copa1),cilindrico(copa1)
cóncavo(copa1),azul(copa1)
ligero(copa1)

Escribe los pasos para obtener la descripción estructural.

8) Describe un dominio siguiendo la misma forma que el ejercicio 7 y plantea una solución para obtener la descripción estructural.

9) Resuelve los problemas 3 y 4 de ID3 usando Weka

10) Resuelve el problema 6 de clustering usando Weka