



Dpto. Lenguajes y Ciencias de la Computación
E.T.S.I. en Informática, Universidad de Málaga

Aprendizaje Computacional

17 de Enero de 2018

Apellidos:
DNI:

Nombre:

PARTE 2

1. Considerando la tabla que se muestra a continuación:

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No

- a) Realiza un árbol de decisión (tipo ID3). Itera hasta el nivel 1. Dibuja el árbol, describe las operaciones y muestra los datos usados en cada una de las iteraciones. (1.5 puntos)
- b) Considera que en el siguiente nivel se considera un algoritmo de poda tal que se etiqueta el nodo con el clase mayoritaria del atributo objetivo (**Play Golf**). Completa el árbol. (0.5 puntos)
- c) Deduce las reglas de decisión del árbol y clasifica las dos primeros y las dos últimas entradas del Dataset indicando si se jugara o no al golf. (0.5 puntos).
2. En el siguiente dataset (el fichero lo puedes encontrar en el campus virtual) tenemos un serie de votaciones de congresistas de EEUU. La estructura del fichero es la siguiente:

handicapped infants: n, y, u
water project cost sharing: n, y, u
adoption of the budget resolution: n, y, u
physician fee freeze: n, y, u
el salvador aid: n, y, u
religious groups in schools: n, y, u
anti satellite test ban: n, y, u
aid to nicaraguan contras: n, y, u
mx missile: n, y, u
immigration: n, y, u
synfuels corporation cutback: n, y, u
education spending: n, y, u
superfund right to sue: n, y, u
crime: n, y, u
duty free exports: n, y, u
export administration act south africa: n, y, u



Dpto. Lenguajes y Ciencias de la Computación
E.T.S.I. en Informática, Universidad de Málaga

Aprendizaje Computacional

17 de Enero de 2018

Apellidos:
DNI:

Nombre:

votation: democrat, republican

El atributo de clasificación es **votation** que clasifica el voto en republicano (**republican**) o demócrata (**democrat**). El resto de atributos reflejan las opiniones del votante en diferentes cuestiones (como por ejemplo, ayudar al Salvador **-el salvador aid-**) y pueden estar a favor (**y**), en contra (**n**) o no tener opinión (**u**).

Entrena un árbol Rpart con el conjunto de datos anterior (eliminando la variable **physician fee freeze**)
Dibuja el árbol y muestra las reglas de clasificación. Imprime la tabla CP (Parámetro de complejidad) y determina usando R el valor apropiado para el parámetro CP y úsalo para podar el árbol. Dibuja el árbol podado. (1 puntos)

Predice los resultados con el árbol podado (el fichero lo puedes encontrar en el campus virtual). Imprime el accuracy de ambos arboles ¿Puede existir un árbol podado con mejor accuracy al que hemos encontrado? Justifica tu respuesta. Si es posible encuentra tal árbol ; imprimelo y muestra su accuracy. (1 puntos)

Explica que ocurre y porque cuando se añade el campo **physician fee freeze**. (0.5 puntos)

4. Teniendo en cuenta los siguientes conjuntos de datos:

A)

X1	X2	Y
1	1	-1
3	3	1
1	3	1
3	1	-1
2	2.5	1
3	2.5	-1
4	3	-1

B)

X1	X2	Y
1	1	-1
3	3	1
1	3	1
3	1	-1
2	2.5	1
3	2.5	-1
4	3	-1
1.5	1.5	1
1	2	-1

Determina si son separables linealmente e indica cual seria la función Kernel mas adecuada en ambos casos (indica tipo de función y sus parámetros) (0.5 puntos).

Sugerencia: Pinta ambos conjuntos de puntos.

Calcula los siguientes parámetros de la Maquina de Soporte Vectorial que podemos obtener con los dataset anteriores y el Kernel elegido:



Dpto. Lenguajes y Ciencias de la Computación
E.T.S.I. en Informática, Universidad de Málaga

Aprendizaje Computacional

17 de Enero de 2018

Apellidos:
DNI:

Nombre:

- Vectores Soporte. (0.25 puntos)
- Ancho del canal (0.25 puntos)
- Vector de Pesos normal al Hiperplano (W) (0.25 puntos)
- Vector B (0.25 puntos)
- La ecuación del Hiperplano y de los planos de soporte positivo y negativo. (0.25 puntos)
- Determinar la clase a la que pertenece un punto dado. (0.25 puntos)
- Pinta el conjuntos de puntos y el Hiperplano. (0.25 puntos)
- Clasifica los puntos (4, 2.5) y (4,1). (0.25 puntos)