TEMA 6 ÁRBOLES DE DECISIÓN

Ganancia de información. Qué expresión correcta, corresponde a la ganancia de y condicionado a x: {

= *IG(Y | X) = E(Y) – E(Y | X).*

~ *IG(Y | X) = E(Y) – E(X).*

~ *IG(Y | X) = E(X) + E(Y | X).*

}

Explicación: IG(Y | X) = E(Y) – E(Y | X), la ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.

Entropía: Hablamos de "DISTRIBUCIÓN EN PICO" cuando se cumple que:{

= Pi = 1 y Pj=0, para todo j≠i

~ Pi = 0 y Pj=1, para todo j≠i

~ Pi=1 y Pj=0, para todo j=i

}

Explicación: Como bien viene indicado en el tema 6 página 5, la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información. O lo que es lo mismo Pi = 1 y Pj=0, para todo j≠i

¿Qué es la entropía?:{

~ Es la medida del grado de incertidumbre asociado a una distribución uniforme.

~ Es la medida del grado de incertidumbre siempre que consiga la máxima información.

= Es la medida del grado de incertidumbre asociado a una distribución de probabilidad.

}

Explicación: La tercera opción es la correcta. La distribución uniforme o de pico (en esta última es en la que conseguimos máxima información) son tipos de distribuciones. Sin embargo, la entropía es una medida y trabaja con distribuciones de probabilidad.

Los árboles de decisión {

~ Es interesante aprenderlos a partir de un grafo

~ Establecen el vector de entrada.

= Ninguna de las anteriores

}

Explicación: La solución es la c) porque es interesante aprenderlos a través de un conjunto de vectores(no grafos) y establecen el orden para testar los atributos y conseguir la clasificación del vector de entrada.

En cuanto a la entropía:{

~ Es mínima cuando la distribución es uniforme.

~ Es máxima cuando la distribución es de pico.

= Es máxima cuando la distribución es uniforme.

}

Respuesta: en una distribución uniforme, todos los valores son igualmente probables Pi=1/N y por tanto la entropía es máxima, lo cual indica máxima incertidumbre (transparencia 5)

Los arboles de decisión establecen {

= en que orden testar los atributos para conseguir la clasificación del vector de entrada.

~ a que velocidad testar los atributos para conseguir la clasificación del vector de salida.

~ el estado inicial de la clasificación de los nodos.

}

Explicación: Según los apuntes, en la página 3 del Tema 6. Una de las características de los árboles de decisión es que: establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada

¿Qué es la ganancia de información? {

~ Conjunto de todos los valores posibles de X

= Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.

~ Ninguna de las anteriores

}

Explicación: La ganancia de información es la medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.

Sobre atributos con gran número de valores: {

~ Se forman grupos pequeños de ejemplos que siempre deben ser homogéneos.

= Se forman grupos pequeños de ejemplos que pueden ser homogéneos por casualidad.

~ Ningunas de las anteriores.

}

Explicación: Transparencia 11 del Tema 6. Se observa que en el apartado de atributos con gran número de valores dice que se forman grupos pequeños de ejemplos que pueden ser homogéneos por casualidad.

Cuando el conocimiento de X mejora la información que se dispone sobre Y, se dice que tiene un entropía condicionada :{

= menor que E(Y).

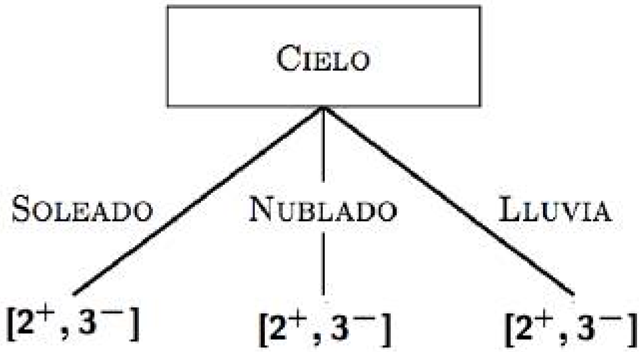
~ igual que E(Y).

~ mayor que E(Y).

}

Explicación: en el tema 6 diapositiva 6. Explica que una entropía condicionada menor que E(Y) indica que el conocimiento de X mejora la información que se dispone sobre Y.

¿Cuál será la entropía del siguiente atributo?



{

= 1.585

~ Para atributos uniformemente distribuidos la entropía será máxima, es decir, 1

~ Para atributos uniformemente distribuidos la entropía será mínima, es decir, 0

}

Explicación: En este caso estamos en una situación de una variable uniformemente distribuida por lo que su entropía será máxima. Sin embargo, según la ecuación para calcular la entropía: E(Cielo) = -(1/3)log2(1/3) - (1/3)log2(1/3) - (1/3)log2(1/3) = 1.585

¿ID3 con qué tipo de atributos trabaja? {

~ Atributos discretos

= Todas las anteriores

~ Atributos continuos si están descompuestos en rangos. }

Explicación: Ambas respuestas son correctas, ID3 trabaja con atributos discretos y, si utilizamos continuos hay que descomponerlos en rangos ordenando los ejemplos según el valor y tomando como límites los puntos medios de aquellos en que se cambie la clase.

En una distribución pico en la que Pi = 1 y Pj=0, para todo j≠i la entropía será{

~ Máxima

= Mínima

~ No lo podemos saber }

Explicación: Tal y como aparece en la transparencia número 5 del tema, bajo estas condiciones, la entropía es mínima, lo cual indica mínima incertidumbre o sea máxima información.

¿Qué implica una alta ganancia de información? {

~ Que el atributo X permite aumentar la incertidumbre de la clasificación del ejemplo de entrada.

= Que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.

~ Que la clase Y permite reducir la incertidumbre de la clasificación del ejemplo de entrada.

}

Explicación: Por definición, una alta ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.

¿Cual de las siguientes afirmaciones es cierta acerca de la entropía?:{

= En una distribución uniforme, todos los valores son equiprobables, lo que implica entropía máxima y por tanto, total incertidumbre.

~ En una distribución pico en la que Pi = 0 y Pj = 1 para j != i la entropía es mínima.

~ En una distribución uniforme todos los valores son equiprobables, lo que implica una entropía mínima y por tanto, la mínima incertidumbre.

}

Explicación: Las distribuciones uniformes deben su nombre a la equiprobabilidad de sus valores, y por definición, esto implica la máxima entropía y máxima incertidumbre.

Por otro lado, las distribuciones pico cumplen la definición si Pi = 1 y Pj = 0.

El recorrido de un árbol de decisión equivale a una:{

= Búsqueda irrevocable.

~ Búsqueda tentativa.

~ Ninguna de las anteriores.

} Explicación: Los árboles de decisiones se recorren mediante un método voraz, por lo tanto irrevocable

Una de las características de los árboles de decisión es: {

= Establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada.

~ En una distribución uniforme, todos los valores son igualmente probables.

~ Ninguna de las anteriores.

}

Explicación: Características de arboles de decisión:

- Estructura para clasificación de vectores de atributos.

- Establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada.

- Para componer dicho orden se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

- Es interesante aprenderlos a partir de un conjunto de vectores

En la entropía, ¿qué grado de incertidumbre tendremos en una distribución uniforme? {

~ Mínima

= Máxima.

~ No se puede saber

}

Explicación: Como hemos visto en clase, la entropía no es una probabilidad, es una incertidumbre; y en este caso, la incertidumbre es máxima.

El algoritmo ID3, a la hora de descomponer los atributos continuos en rangos, después de ordenarlos: {

= Debemos tener en cuenta los puntos límites donde se cambie de clase.
~ No es necesario descomponer, pues siempre tratamos con atributos discretos.
~ Debemos de realizar la descomposición siempre en partes iguales.
}

Explicación: En primer lugar realizamos la ordenación de los ejemplos, y luego, realizamos la descomposición tomado como referencia los puntos medios de aquellos que se cambien de clase.

Lanzamos un dado al aire, ¿cuál es la entropía del lanzamiento?:{

- 1, máxima incertidumbre

- 0,16, al ser distribución uniforme

= 2,58, máxima incertidumbre.

}

Explicación: la entropía se calcularía así: 6*(-(1/6)*log(1/6)/log2)

Es distribución uniforme debido a que tenemos la misma probabilidad de que nos toque una de las 6 caras, por eso multiplicamos por 6.

Una entropía condicionada menor que E(Y) indica que: {

~ El conocimiento de X reduce la información que se dispone sobre Y.

= El conocimiento de X mejora la información que se dispone sobre Y.

~ No indica nada. Siempre es mayor o igual que E(Y).

}

Explicación: La entropía condicionada puede ser menor que E(Y), lo cual indica que hay una mejora en la información que se tiene de Y.

En la Entropía condicionada: {

~ No interviene ninguna variable.

~ Interviene una variable.

= Intervienen al menos dos variables.

}

Explicación: La entropía condicionada es una extensión del concepto de entropía de la información a procesos donde intervienen varias variables aleatorias no necesariamente independientes.

La Entropía condicionada es según su definición :{

~ la entropía de la distribución de Y condicionada a Y.

= la entropía de la distribución de Y condicionada a X.

~ la entropía de la distribución de X condicionada a X.

}

Explicación: en el tema 6 diapositiva 6. Nos dice en la definición de entropía condicionada que es la entropía de la distribución de Y condiciona a X.

¿Qué pasa si usamos ID3 con atributos continuos?{

~ Nada porque es con lo que trabaja.

= Hay que descomponerlo en rangos.

~ Hay que ordenarlos según valor y pasarlos a discretos.

}

Explicación: ID3 solo trabaja con atributos discretos, para el caso de utilizar atributos continuos hay que descomponerlo en rangos. La primera esta claro que esta mal, pero la tercera esta mal porque dice "Hay que ordenarlos según valor" (hasta ahí esta bien) "y pasarlos a discretos" (esto ya esta mal).

Los árboles de decisión, ¿contienen información local? {

~ No, no están basados en información local.

= Sí, lo que permite poder encontrar la mejor solución.

~ Si, pero nunca encontrarán la mejor solución.

}

Explicación: Los árboles de decisión usan una estrategia irrevocable de descenso por gradiente, lo que quiere decir que poseen conocimiento local y buscan la mejor solución a partir de esa información local.

El valor de entropía, ¿está acotado de 0 a 1?: {

~ Si, dado que es un valor de probabilidad.

~ Si, dado que es una medida del grado de incertidumbre.

= No, dado que es una medida del grado de incertidumbre.

}

Explicación: Al no ser un valor de probabilidad, cuanto mas cercano a 0 sea el valor menos incertidumbre existe pero el valor puede ser mayor que 1, es decir, 1 no quiere decir que el valor de incertidumbre sea máximo.

Si tenemos una E(Y) = 0.971 y tras analizar la información de la que disponemos obtenemos los siguientes datos para aplicar una entropía condicionada:

vj	**Prob(X= vj)**	**E(Y|X = vj)**
Atributo1	**0.3**	**0.92**
Atributo2	**0.4**	**0.81**
Atributo3	**0.3**	**0.92**

¿Después de calcular el valor de E(Y | X) podemos decir que hemos obtenido ganancia de información?

{

= Si, hemos obtenido una ganancia ≈ 0.09

~ Si, hemos obtenido una ganancia ≈ 0.5

~ No

}

Explicación: *He utilizado los datos de la diapositiva 9, por si no queda clara la explicación.*

** Primero aplicamos la formula de la entropía condicionada a los datos de la tabla.*

E(Y | X) = Σj Prob(X= vj) E(Y | X = vj)

E(Y | X) = 0.3 * 0.92 + 0.4 * 0.81 + 0.3 * 0.92 = 0.876

** Una vez hecho esto aplicamos la de la ganancia*

IG(Y | X) = E(Y) – E(Y | X)

IG(Y | X) = 0.971 - 0.876 = **0.095 ≈ 0.09**

La fórmula de la entropía condicionada es:{

= E(Y | X) = Σj Prob(X= vj) E(Y | X = vj)

~ E(X | Y) = Σj Prob(X= vj) E(Y | X = vj)

~ E(Y | X) = Σj Prob(Y= vj) E(Y | X = vj)

}

Explicación: La fórmula correcta para el cálculo de la entropía condicionada, como pone en la transparencia 6 del tema 6, es E(Y | X) = Σj Prob(X= vj) E(Y | X = vj)

Dado un conjunto de ejemplos y su árbol de decisión generado por ID3. Los elementos de dicho árbol son:{

~ Nodos que contienen valores del padre, arcos que contienen variables continuas y hojas que clasifican el ejemplo como positivo.

~ Nodos que contienen valores del padre, arcos que contienen valores del hijo y las hojas que clasifican el ejemplo como positivo.

= Nodos que contienen atributos, arcos que contienen posibles valores del padre y las hojas que clasifican el ejemplo como positivo o negativo.

}

Explicación: Cada nodo del árbol representa cada uno de los atributos (Antigüedad, Moroso, Ingresos). Los arcos representan los valores que el nodo puede tomar (<600, 600-1200, >1200) y las hojas representan la decisión final que pueden ser positivos o negativos (Conceder, No conceder).

En los árboles de decisión, que criterio utilizamos para el orden en el que testaremos los atributos: {

~El que nos da una ganancia de información más baja.

~Es indiferente, el orden es aleatorio.

=El que nos da una ganancia de información más alta.

}

Explicación: Para componer dicho orden se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

Ej.	Cielo	Humedad	JugarTenis
D1	Soleado	Alta	-
D2	Soleado	Alta	-
D3	Nublado	Alta	+
D4	Lluvia	Alta	+
D5	Lluvia	Normal	+
D6	Lluvia	Normal	-
D7	Nublado	Normal	+
D8	Soleado	Alta	-
D9	Soleado	Normal	+
D10	Lluvia	Normal	+
D11	Soleado	Normal	+
D12	Nublado	Alta	+
D13	Nublado	Normal	+
D14	Lluvia	Alta	-

Dado el conjunto anterior, que atributo cogeríamos primero para aprender el concepto “días que se juega a tenis” y obtener el nodo inicial del árbol de decisión mediante el algoritmo ID3{

~ El orden en que cojamos los atributos no tiene importancia, el nodo inicial puede ser tanto “cielo” como “humedad”.

= Cogemos el atributo “cielo”, ya que es el que mayor ganancia de información nos ofrece.

~ Cogemos el atributo “humedad”, ya que es el que mayor ganancia de información nos ofrece.

}

Explicación: Para escoger un nodo del árbol a construir, hay que calcular la ganancia de información que nos proporciona cada atributo. Una vez calculadas, Escogemos el atributo que tenga mayor ganancia de información:

Las fórmulas para calcular la Entropía, la Entropía Condicionada y la Ganancia de Información las encontramos en las transparencias 5, 6 y 7, respectivamente.

E(Jugar) = $-(9/14)*\log_2(9/14) - (5/14)*\log_2(5/14) = 0.94$

E(Jugar | Cielo) = Prob(Nublado)*E(Jugar | Nublado) + Prob(Soleado)*E(Jugar | Soleado) + Prob(Lluvia)*E(Jugar | Lluvia)

E(Jugar | Cielo) = $(4/14 * 0) + (5/14 * (- (3/5) * \log_2(3/5) – (2/5) * \log_2(3/5)) + (5/14 * (- (3/5) * \log_2(3/5) – (2/5) * \log_2(3/5)) = 0.694$

IG (Jugar | Cielo) = E(Jugar) - E(Jugar | Cielo) =0.94 – 0.694 = 0.246

E(Jugar | Humedad) = Prob(Alta)*E(Jugar | Alta) + Prob(Normal)*E(Jugar |

Normal)

E(Jugar | Humedad) = $(7/14 * (- (4/7) * \log_2(4/7) – (3/4) * \log_2(3/4)) + (7/14 * (- (1/7) * \log_2(1/7) – (6/7) * \log_2(6/7)) = 0.789$

IG (Jugar | Humedad) = E(Jugar) - E(Jugar | Humedad) =0.94 – 0.79 = 0.151

POR LO TANTO EL PRIMER NODO DE NUESTRO ÁRBOL SERÍA EL ATRIBUTO "CIELO"

Si tenemos la mínima incertidumbre de que el equipo A ganará al equipo B porque el partido está amañado podemos afirmar que{

~ Podemos apostar por A, pero no sabemos nada con certeza

= Tenemos máxima información sobre ese partido y podremos apostar seguros por A

~ Deberemos apostar por B; porque seguro que ganara

}

Explicación: Tener la mínima incertidumbre (0) significa que tenemos la máxima información sobre ese caso; y no se puede dar el caso contrario bajo ningún motivo. Por eso es seguro apostar que el equipo A ganará.

¿Que queremos conseguir con un árbol de decisión?

{

~Explorar las infinitas soluciones de un grafo y dar una aleatoria.

~Hacer un grafo de 20 nodos.

=Saber decidir si tengo que realizar una acción o no.

}

Explicación: En un árbol de decisión se seleccionan sólo aquellos atributos que nos ofrecen mejor ganancia de información, no todos. Los arboles de decisión nos sirven para determinar si una acción/operación se debe realizar o no.

¿A qué se debe que el cálculo de la entropía pueda superar el valor 1?:{

= A que existen más de 2 valores en la distribución de probables.

~ A que existen 2 o menos valores en la distribución de probables.

~ El cálculo de la entropía nunca supera el 1.

}

Explicación: Como vimos en clase, en caso de que tengamos solo 2 posibilidades (el ejemplo "Cara/Cruz de la moneda) tendremos una entropía acotada entre 0 y 1, mientras que a la hora de calcular con más de dos (Ganar/Perder/Empatar...) esta puede rebasar el valor de la unidad.

¿Cuál de las siguientes afirmaciones es correcta?:{

~ En los atributos numéricos el ID3 trabaja con valores discretos y continuos.

= En los atributos numéricos el ID3 trabaja sólo con valores discretos.

~ En los atributos numéricos el ID3 trabaja sólo con valores continuos.

}

Explicación: Como podemos observar en la transparencia 11 el ID3 sólo trabaja con valores discretos. En el caso de que fuesen continuos habría que descomponerlos en rangos.

¿Qué afirmación acerca de la entropía es falsa? {

~ En una distribución uniforme la entropía es máxima.

~ La entropía mide el grado de incertidumbre asociado a una distribución de probabilidad.

= En una distribución pico la información es mínima.

}

Explicación: La falsa es la ultima puesto que en una distribución pico la entropía es mínima por lo que la incertidumbre es mínima pero la información es la máxima.

Los árboles de decisión establecen en qué orden testar los atributos para conseguir la clasificación del vector de entrada. Ahora, ¿cómo se compone dicho orden? {

~ se eligen primero aquellos atributos que peor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

= se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

~ se eligen los atributos sin importar que ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

} Explicación: El orden en el que se testan los atributos de los vectores de entrada para conseguir su clasificación se compone escogiendo primero los atributos que mejor ganancia de información prometen.

La evaluación de un árbol de decisión se realiza: {

= Siempre de derecha a izquierda.

~ Siempre de izquierda a derecha.

~ Ninguna de las anteriores son correctas.

}

Explicación: Las variables cuyo valor se conoce antes de tomar la decisión , han de aparecer a la izquierda del nodo en el desarrollo del árbol; las que no se conocen aparecerán a la derecha.

El Algoritmo ID3 que trabaje con atributos continuos tiene que ...{

= ...descomponerlos en rangos.

~ ...homogeneizar los rangos por casualidad.

~ ...penalizar los atributos de valor elevado.

}

Explicación: "ID3 sólo trabaja con atributos discretos. Si se usan atributos continuos hay que descomponerlos en rangos. " *(Transparencia 11 del Tema 6)*

En un árbol de decisión es cierto que

{

~ Siempre se seleccionan todos los atributos

~Siempre se seleccionan n-1 atributos, siendo n el número de atributos

=No siempre se seleccionan todos los atributos

}

Explicación: En un árbol de decisión se seleccionan sólo aquellos atributos que nos ofrecen mejor ganancia de información, no todos o n-1, como se puede ver en los ejemplos del tema y como se comentó en clase.

¿Cual de las siguientes afirmaciones es cierta?(considerando una entropía de distribución de Y condicionada a X){

~ Una entropía condicionada mayor que E(Y) indica que el conocimiento

de X mejora la información que se dispone sobre Y

~ Una entropía condicionada menor que E(Y) indica que el conocimiento

de X no mejora la información que se dispone sobre Y

= Ambas son falsas

}

Explicación: Transparencia 6:

"Una entropía condicionada menor que E(Y) indica que el conocimiento

de X mejora la información que se dispone sobre Y"

Dado que N es el número de valores que puede tomar una variable, el valor de su entropía se encuentra acotado por: {

= [0, log2 (N)]

~ [0, ∞)

~ No es posible determinarlo/ninguna de las anteriores

}

Explicación:

En el caso de una distribución pico, la entropía es igual a cero (pag. 5 de las transparencias) y, en el caso de una distribución uniforme, los valores de todas las probabilidades son 1/N.

Por lo tanto la entropía es la sumatoria de -1/N*log2 (1/N), dado que todos los términos son iguales, sustituimos la sumatoria por una multiplicación:

N (-1/N*log2 (1/N)) = -log2 (1/N) = - (log2 (1) - log2 (N)) = - (0 - log2 (N)) = log2(N)

NOTA: Tal vez sería mas adecuado colocar como respuesta correcta el intervalo [0, -log2 (1/N)], para que no fuese necesario aplicar las identidades logarítmicas.

Respecto a la entropía… {

~ En una distribución uniforme, todos los valores son igualmente probables Pi = 1/N y por tanto la entropía es mínima, lo cual indica máxima incertidumbre.
~ Es una medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el valor de otra Y.
= Ninguna de las anteriores es correcta.

}

Explicación: La entropía es una medida del grado de incertidumbre asociado a una distribución de probabilidad, en una distribución uniforme, todos los valores son igualmente probables Pi = 1/N y por tanto la entropía es máxima, lo cual indica máxima incertidumbre, por el contrario, en una distribución pico en la que Pi = 1 y Pj = 0, para todo j != i la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

Sabiendo que el resultado de los partidos disputados entre el Hércules y el Elche ha sido:

Gana Hércules(H): 20

Empate (X): 5

Gana Elche(E): 5

Calcula la entropía de que el Hércules gane al Elche en un partido de futbol. {

~ E(H) = 0

= E(H) = 1.25

~ E(H) =0.5

}

Explicación: E(H) = -(2/3)(log(2/3)/(log2) - 2 * ((1/6)log(1/6)/(log2)) = 1.25

La fórmula que nos permite calcular la Ganancia Normalizada es: {

$$G_N(S,A) = \frac{G(S,A)}{\sum v_i \in V(A) - p_{v_i} \log_2 p_{v_i}}$$

=

$$G_N(S,A) = \frac{\sum v_i \in V(A)}{G(S,A) - p_{v_i} \log_2 p_{v_i}}$$

~

$$G_N(S,A) = \frac{G(S,A)}{p_{v_i} \log_2 p_{v_i} - \sum v_i \in V(A)}$$

~

}

Explicación: La fórmula para calcular la ganancia normalizada es:

$$G_N(S,A) = \frac{G(S,A)}{\sum v_i \in V(A) - p_{v_i} \log_2 p_{v_i}}$$

Referente a los árboles de decisión.... :{

~ es interesante aprenderlos a partir de un conjunto de matrices

= es interesante aprenderlos a partir de un conjunto de vectores

~ ninguna de las anteriores.

}

Explicación: T6, trasparencia 3. Sirven para la clasificación de vectores de atributos (vector de entrada).

ID3 sólo trabaja con atributos discretos. Para tratar atributos numéricos o continuos hay que…{

- ignorarlos

= descomponerlos en rangos

- descomponerlos como máximo en tres valores discretos

}

Explicacion:

Si se usan atributos continuos hay que descomponerlos en rangos. Para ello se ordenan los ejemplos según el valor y se toman como puntos límite los puntos medios de aquellos en que se cambie de clase. (T6, pag.11)

La complejidad de ID3: {

~ ID3 crece linealmente con el número de ejemplos de entrenamiento.

~ ID3 crece exponencialmente con el número de atributos.

= Ambas son correctas.

}

Explicación: El espacio ocupado por id3 es el ocupado por el árbol de decisión. En el peor de los casos, habrá un nodo hoja por cada ejemplo, con un número de nodos intermedios igual al tamaño del espacio de instancias. Esto se produciría por ejemplo, cuando se le pasara como entrada todos los posibles ejemplos, cada uno con una clase diferente. Como intentaría encontrar una forma de separar todos los ejemplos, esto le llevaría a considerar todos las posibles combinaciones de valores de atributos (tamaño de espacio de instancias). Evidentemente, esto no es un caso real. En el mejor de los casos (por ejemplo, todos los ejemplos pertenecen a la misma clase), habrá un único nodo.

Se lanza una moneda al aire para ver si sale cara o cruz (dos estados con probabilidad 0,5). Su entropía es: {

~ 0,5

~ 0

= 1

}

Explicación: aplicando la fórmula de la entropía, $E = 0{,}5 \cdot \log_2(1/0{,}5) + 0{,}5 \cdot \log_2(1/0{,}5) = (0{,}5+0{,}5) \cdot \log_2 2 = \mathbf{1}$

Calcular el valor de Ent ([5,2,1]) {

= 1.3

~ 1.4

~ 1.5

}Explicación: Ent([5,2,1]) = -5/8 log2 (5/8) - 1/4 log2 (1/4) - 1/8 log2 (1/8)

Según la definición de Ganancia de Información . ¿Cuál de estas respuestas es correcta? : {

~ Ganancia de información es conocer el valor de una variable aleatoria X.

~ Si obtenemos una alta ganancia implica que el atributo X permite ampliar la incertidumbre de la clasificación.

= Una alta ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada. }

Explicación : Ganancia de información es la medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y. Una alta ganancia implica que el atributo X permite reducir la incertidumbre de la clasificación del ejemplo de entrada.

En los árboles de decisión: {

= Se usan algoritmos voraces.

~ Vuelta atrás, (Backtracking).

~ Ambas son correctas. }

Explicación: No permite la vuelta atrás en los atributos seleccionados. Utiliza el esquema voraz, elegir uno y filtrar recursivamente.

Si en un dado de 4 caras en el que todas las cara tiene la misma posibilidad de aparecer eliminamos dos caras(nunca pueden salir) que tipo de incertidumbre existiría?

{

=máxima.

~mínima.

~ninguna de las anteriores.

}

Explicación: Aunque se eliminen valores de la probabilidad si estos son igual de probable la incertidumbre sigue siendo máxima.

Sobre la entropía y el grado de incertidumbre sobre una distribución de probabilidad, ¿Cuál es la opción correcta?:{

~ La entropía es mínima cuando existe una máxima incertidumbre.

~ Cuando una distribución de probabilidad es uniforme la entropía es mínima.

= La entropía es máxima cuando existe una máxima incertidumbre.

}

Explicación: La entropía es máxima cuando existe una distribución de probabilidad uniforme (Pi = 1/N) o máxima incertidumbre. Es mínima cuando existe una distribución pico (Pi = 1 y Pj = 0, para todo j≠i) o mínima incertidumbre.

En el algoritmo ID3: {

~ Se calcula la ganancia de información de los atributos más prometedores.

~ Se escoge el atributo con menor entropía o grado de incertidumbre.

= *Ninguna de las anteriores.*

}

Explicación: se calcula la ganancia de información de todos los atributos y se escoge el de mayor resultado.

En una distribución uniforme…:{

~ La entropía es máxima, lo cual indica mínima incertidumbre.

~ La entropía es mínima, lo cual indica máxima información.

= La entropía es máxima, lo cual indica mínima información.

}

Explicación: En una distribución uniforme, todos los valores son igualmente probables Pi = 1/N y por tanto la entropía es máxima, lo cual indica máxima incertidumbre (mínima información).

¿Cuál de estas afirmaciones es falsa?

{

~Medida del grado de incertidumbre asociado a una distribución de probabilidad.

=La entropía te da el grado de probabilidad

~La entropía es máxima indica que hay mínima información.

}

Explicación: La entropía te da un grado de incertidumbre, en vez de uno de probabilidad.

Sobre la entropía en el contexto de IA, se puede decir:{

= Es la medida del grado de incertidumbre asociado a una distribución de probabilidad.

~ Es interesante aprenderla a partir de un conjunto de vectores.

~ Es siempre una distribución uniforme.

}

Explicación: La utilización de vectores es normalmente usada para lo estudio de árbol de decisiones. La entropía puede ser una distribución uniforme, pero ni siempre lo es. Luego solamente la primera opción es correcta.

Deseamos generar un árbol de decisión para saber si un terreno es apto para viñedo. Para ello partimos de los atributos y valores de la siguiente tabla:

VARIABLES (Viñedo)					
Casos	Lluvia	Temperatura	Humedad	Fertilidad	Si/No
V1	Alta	Irregular	Alta	Normal	No
V2	Alta	Irregular	Normal	Alta	No
V3	Media	Irregular	Alta	Normal	No
V4	Baja	Regular	Alta	Normal	Si
V5	Baja	Regular	Alta	Normal	Si
V6	Baja	Regular	Normal	Alta	Si
V7	Media	Regular	Normal	Alta	Si
V8	Alta	Regular	Alta	Normal	No
V9	Alta	Regular	Alta	Normal	No
V10	Baja	Regular	Normal	Alta	Si
V11	Alta	Regular	Alta	Alta	No
V12	Media	Irregular	Alta	Alta	Si
V13	Media	Irregular	Normal	Normal	No
V14	Baja	Irregular	Normal	Normal	No
Totales	Si =	6	No =	8	

¿Cuál sería el primer atributo del árbol?

{

~ Temperatura.

~ Fertilidad.

= Lluvia.

}

Explicación: En el siguiente cuadro se pueden observar los cálculos efectuados y vemos que la ganancia del atributo lluvia es mayor que la de cualquier otro.

Entropía Inicial (6+, 8-) =			0,9852			
Valores		Veces	+	-	Entropía	Ganancia
Lluvia	Alta	5	0	5	0,0000	0,4417
	Media	4	2	2	1,0000	
	Baja	5	4	1	0,7219	
	Total/prob.	14	3/7	4/7		
Temperatura	Irregular	6	1	5	0,6500	0,1613
	Regular	8	5	3	0,9544	
	Total/prob.	14	3/7	4/7		
Humedad	Alta	8	3	5	0,9544	0,0113
	Normal	6	3	3	1,0000	
	Total/prob.	14	3/7	4/7		
Fertilidad	Alta	6	4	2	0,9183	0,1281
	Normal	8	2	6	0,8113	
	Total/prob.	14	3/7	4/7		

El árbol de decisión es utilizado en problemas de... {

~clasificación

~inducción

=ambos son correctos }

Explicación: Los árboles de decisión toman como entrada una situación descrita por un conjunto de atributos y devuelve una decisión que es el predizido valor al valor de entrada.

Esperamos un mensaje que puede consistir de las letras en minúscula de la a hasta la z.

Cual de las siguientes afirmaciones es correcta cuando recibamos el mensaje "qalmnbphijcdgketrsfuvxyzwño"?

{

~ Es una distribución pico en la que Pi = 1 y Pj=0, para todo j≠i la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información

= Es una distribución uniforme, todos los valores son igualmente probables Pi = 1/N y por tanto la entropía es máxima, lo cual indica máxima incertidumbre

~ Ninguna de las anteriores es correcta

} Explicación:

Se puede decir que este mensaje llega a nosotros con la máxima entropía (o desorden posible); ya que es poco probable que se pueda pronosticar la entrada de caracteres, pues estos no se repiten ni están ordenados en una forma predecible.

¿Cuál de las siguientes características de los árboles de decisión es incorrecta? {

~ Es una estructura para la clasificación de vectores de atributos.

~ Establece en qué orden testear los atributos para conseguir la clasificación del vector de entrada.

= Se eligen primero aquellos atributos que menor ganancia de información.

}

Explicación: En los árboles de decisión se eligen primero aquellos atributos con mayor ganancia de información para obtener antes una resolución del problema.

Atendiendo a los datos de la siguiente tabla:

X	Y
Moneda Azul	**Cara**
Moneda Roja	**Cara**
Moneda Azul	**Cruz**
Moneda Verde	**Cara**
Moneda Azul	**Cruz**
Moneda Roja	**Cara**
Moneda Roja	**Cruz**
Moneda Verde	**Cruz**

La Ganancia de Información IG(Y|X) sería {

= IG(Y|X) = 0.15625

~ IG(Y|X) = 0

~ IG(Y|X) = 0.84375

}

Explicación:

Al agrupar la información que nos dan obtenemos la siguiente tabla:

v_j	Prob ($X = v_j$)	E($Y\|X = v_j$)
Moneda Azul	3/8	0.875
Moneda Roja	3/8	0.875
Moneda Verde	1/4	0.75

Donde obtemos las entropías para los tres posibles valores de v_j:

E(Y|X = Moneda Azul) = $-1/4\cdot\log_2(1/4)-1/8\cdot\log_2(1/8) = 0.875$

E(Y|X = Moneda Roja) = $-1/8\cdot\log_2(1/8)-1/4\cdot\log_2(1/4) = 0.875$

E(Y|X = Moneda Verde) = $-1/8 \cdot \log_2(1/8) - 1/8 \cdot \log_2(1/8) = 0.75$

Por lo tanto para obtener la Entropía Condicionada no hay más que sumar las multiplicaciones de la probabilidad de cada variable por su entropía:

E(Y|X) = 3/8*0.875 + 3/8*0.875 + 2/8*0.75 = 0.84375

Con lo que ya podemos obtener la Ganancia de Información, previamente calculando la Entropía de Y:

$E(Y) = -1/2 \cdot \log_2(1/2) - 1/2 \cdot \log_2(1/2) = 1$

IG(Y|X) = 1 - 0.84375 = 0.15625

Se nos plantea resolver un problema de arboles de decisiones mediante el algoritmo ID3, y en la especificación del conjuntos de valores nos encontramos con la siguiente tabla adjunta:

ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Cantidad	**200**	**150**	**100**	**350**	**500**	**250**	**400**	**450**	**300**	**550**
Salida	**1**	**1**	**1**	**0**	**1**	**1**	**0**	**1**	**0**	**1**

Elije la respuesta incorrecta: {

~ Habría que ordenar los Ids según la cantidad y tomar como puntos límite los puntos medios de aquellos en que cambie el valor de la salida.

~ Estamos ante un caso de atributos numéricos continuos; por lo que es necesario discretizar estos y descomponerlos en rangos.

= El algoritmo ID3 trabaja con todo tipo de variables, porque el propio algoritmo trata esto sin necesidad de extensiones adicionales.

}

Explicación: En arboles de decisión trabajan con atributos categóricos y el propio algoritmo ID3 también. En caso de problemas con atributos continuos existen extensiones para este, que tratan precisamente de crear rangos para identificar estos valores y tenerlos así distribuidos de manera categórica.

¿Cual de los siguientes es un algoritmo utilizado en Arboles de Decisión?: {

~ Backtracking

~ Ramificación y poda.

= El algoritmo ID3.

}

Explicación: El algoritmo ID3 es utilizado en los arboles de decisión, a partir de un conjunto de ejemplos genera unas reglas.

¿Cuál de las siguientes características corresponde a un árbol de decisión?:

{

= Estructura para clasificación de vectores de atributos.

~ Establece en qué orden testar los atributos para conseguir la clasificación del vector de salida.

~ Se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de salida.

}

Explicación: En las 2 respuestas incorrectas se hace referencia al vector de salida y no al de entrada. Se puede encontrar en la página 3 de las transparencias del tema 6.

La entropía puede ser definida como:{

~Estructura para clasificación de vectores de atributos. ~Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y. = Medida del grado de incertidumbre asociado a una distribución de probabilidad.

}

Explicación: la primera opción se refiere a árboles de decisión y la segunda se refiere ganancia de información. La ultima opción es la correcta.

¿Que es la ganancia de información?:{

~ Estructura para clasificación de vectores de atributos.

= Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.

~ Medida del grado de incertidumbre asociado a una distribución de probabilidad.

}

Explicación: La ganancia de información esta definida de esta manera en la transparencia 7 del tema 6

Respecto a las variables en la clasificación de un arbol de decisión, indica que afirmación es correcta:{

~ Las variables continuas tienen el mismo tratamiento que las variables discretas.

= Las variables continuas deben establecerse dentro de categorías.

~ Ninguna de las anteriores es correcta.

}

Explicación: Las variables continuas deben categorizarse antes de poder ser usadas en un árbol de decisión (deben ser "discretizadas" aunque posteriormente no se seleccionen sus atributos durante el proceso de decisión).

Indica que afirmación no es correcta:{

~ Ganancia de información es la Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de otra Y.

~ Entropía es la medida del grado de incertidumbre asociado a una distribución de probabilidad.

= Ninguna de las anteriores es correcta.

}

Explicación: Hojas 5 e 7 del Tema 6 contienen las informaciones acerca de GI y Entropía.

En un sistema dado, en el que todas las opciones tengan la misma probabilidad: {

~ La entropía es baja.

= La entropía es alta.

~ No hay entropía. }

Explicación: Al haber la misma probabilidad para todas las opciones tenemos una total incertidumbre de cual puede ser el resultado. Ejemplo en la transparencia 5 del tema 6.

En una distribución pico en la que Pi = 1 y Pj=0... : {

=para todo j≠i la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

~para todo j=i la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

~para todo j≠i la entropía es mínima lo cual indica mínima incertidumbre o sea mínima información.

}

Explicación: En una distribución uniforme, todos los valores son igualmente probables Pi = 1/N y por tanto la entropía es máxima, lo cual indica máxima incertidumbre. Por el contrario, en una distribución pico en la que Pi = 1 y Pj=0, para todo j≠i la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

Según la definición de Entropía cuando obtendremos máxima información:{

=Cuando hay mínima incertidumbre

~Cuando la entropía es máxima

~Nunca podemos obtener máxima información

}

Explicación: En una distribución pico en la que Pi = 1 y Pj=0, para todo j≠i la entropía es mínima lo cual indica mínima incertidumbre o sea máxima información.

En cuanto a los árboles de decisión:{

~ Es una estructura donde se clasifican vectores de atributos. Estos atributos se testan aleatoriamente.

~ Se clasifican vectores de atributos testados en orden. Para componer dicho orden se eligen primero los atributos menos probabilidad tienen de llegar a una solución optima.

= Se clasifican vectores de atributos testados en orden. Para componer dicho orden se eligen primero los atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

}

Explicación: Las características de los árboles de decisión son:

Estructura para clasificación de vectores de atributos.

Establece en qué orden testar los atributos para conseguir la clasificación del vector de entrada.

Para componer dicho orden se eligen primero aquellos atributos que mejor ganancia de información prometen a efectos de descubrir la clase del vector de entrada.

Es interesante aprenderlos a partir de un conjunto de vectores

Los árboles de decisión son: {

~ Datos sin ningún tipo de relación entre si.

~ Una estructura para la separación de vectores por tipos de datos.

= Una estructura para clasificación de vectores de atributos.

}

Explicación: Según se muestra en la diapositiva 3 del tema, entre las características de los árboles de decisión, se observa que es una estructura para clasificación de vectores de atributos.

La entropía es máxima cuando: {

~ La incertidumbre es mínima.

~ La ganancia de información es máxima.

= En una distribución los valores son igualmente probables Pi = 1/N y por tanto la entropía es máxima.

}

Explicación: La entropía es una medida de incertidumbre, lo cual es lo contrario a tener información. Es máxima cuando en una distribución uniforme todos los valores son igualmente probables. E(s) = Sumatorio (-p(i) log2(p(i))) = 1.

Se me ha pasado por alto meter una pregunta en el foro en las fechas adecuadas. No ha sido a propósido, pero ha sucedido. Así que me gustaría que me indicase si directamente estoy suspendido. Un saludo y muchas gracias.

Una entropía condicionada menor que E(Y): {

= Indica que el conocimiento de X mejora la información que se dispone sobre Y.

~ Indica que el conocimiento de X reduce la información que se dispone sobre Y.

~ No indica nada.

}

Después de realizar el cálculo de la Entropía condicionada obtenemos que E(Y|X) = 0. ¿Qué podemos determinar? {

~ Nos hemos equivocado en el cálculo.

= El conocimiento de X implica el conocimiento de Y.

~ El conocimiento de X no nos aporta información alguna sobre Y.

}

Explicación: Un valor de E(Y|X) menor que E(Y) mejora la información que se dispone de Y, pero si E(Y|X) = 0 se puede determinar el valor que tendrá Y dado el valor de X.

Tenemos el vector [66,114,66,69,66], ¿cuál es su entropía?{

= 1,37

~ -1,37

~ 0.486

}

Explicación: Aplicando la fórmula de la entropía (pàg 5 T6) tenemos que H = [-3/5 * log (3/5) - 2/5 * log(1/5)]/log2 = 1'37. En cualquier caso la segunda respuesta no podría ser por ser negativa y la última sale de cambiar un signo de la fórmula.

Hallar la ganancia de información con los datos siguientes

X	Y
P1	Yes
P2	Yes
P3	Yes
P2	Yes
P1	Yes
P2	No
P1	Yes
P3	No
P3	Yes
P2	No

:{

= IG(Y|X)=1,243

~ IG(Y|X)=0,104

~ IG(Y|X)=0,578

}

Explicación:

1º Paso: Hay que conocer cuantos “Yes” y cuantos “No” hay por cada Pm.

P1 → 3 (Yes) y 0 (No)

P2 → 2(Yes) y 2 (No)

P3 → 2 (Yes) y 1 (No)

2º Paso: Probabilidad y entropía por cada Pm.

Pm	Prob(X=Pm)	E(Y|X = Pm)
P1	0,3	0
P2	0,4	1
P3	0,3	0,918

Cálculos de la entropía por cada Pm:

P1: -3/3*log2(3/3)-0/3*log2(0/3)=0

P2: -2/4*log2(2/4)-2/4*log2(2/4)=1

P3: -2/3*log2(2/3)-1/3*log2(1/3)=0,918

3º Paso: Entropía condicionada.

E(X|Y)=0,3*0+0,4*1+0,3*0,918=0,675

4º Paso: Suma de las entropías de cada Pm.

E(Y)=0+1+0,918=1,918

Último paso: La ganancia de información.

IG(Y|X)=1,918-0,675=**1,243**

La ganancia de información se puede calcular de la siguiente manera:{

= IG(Y | X) = E(Y) - E(Y |X)

~ IG(X) = E(Y) - E(Y |X)

~ IG(Y | X) = E(X) - E(Y |X)

}

Explicación: La fórmula correcta se puede encontrar en la transparencia 7

Si después de lanzar un dado cargado 100 veces tenemos que cada una de las caras ha salido el siguiente número de veces:

1 -> 8

2 -> 11

3 -> 6

4 -> 7

5 -> 9

6 -> 59

¿Qué entropía tenemos sobre el lanzamiento de dicho dado? {

~ 2.5849

= 1.9156

~ 1.0152

}

Explicación: Calculamos la entropía:

E(dado) = - 8/100*log2(8/100) - 11/100*log2(11/100) - 6/100*log2(6/100) - 7/100*log2(7/100) -

9/100*log2(9/100) - 59/100*log2(59/100) = 1.9156

Como curiosidad añadir que el primer valor (2.5849) es la entropía que producirá un dado sin cargar, en el que cada valor aparecerá 1 vez de cada 6 lanzamientos, de media.

$$a)\ E(S) = \sum_{i \in C}^{n} -p_i log_2 p_i$$

$$b)\ E(S) = \sum_{i \in C}^{n} -p_i + p_i$$

$$c)\ E(S) = \sum_{i \in C}^{n} -(p_i log_2 p_j)^2$$

Teniendo en cuenta las formulas anteriores cual se corresponde al calculo de la Entropía {

= Opción A.

~ Opción B.

~ Opción C.

}

Explicación: Si observamos en la transparencia 6 del Tema 6, observamos cual es la formula correcta que corresponde con la opción 1

¿Qué es la Ganancia de información? :{

= Medida de cuanto ayuda el conocer el valor de una variable aleatoria X para conocer el verdadero valor de *otra Y.*

~ Medida de cuanto ayuda a no conocer el valor de una variable Y para conocer el verdadero valor de *otra Y.*

~ Ninguna de las anteriores.*}*

Explicación: La correcta es la primera opción como podemos ver en la pagina 7 del tema 6