**TEMA 8. APRENDIZAJE BAYESIANO (CON EXPLICACIÓN)**

1. El teorema de Bayes aplicado al aprendizaje permite:{

 = Conocer el máximo a posteriori MAP.

~ Determinar si los ejemplos del conjunto de entrenamiento son condicionalmente independientes.

~ Conocer la probabilidad a priori de las clases de un problema de aprendizaje.

}

Explicación: Ese es el fundamento del teorema de Bayes.

1. Cuántos y qué tipos de aprendizaje podemos encontrarnos en el aprendizaje automático:{

~ 2 tipos: supervisado y por refuerzo.

= 3 tipos: supervisado, no supervisado y por refuerzo.

~ 2 tipos: supervisado y por inducción.

}

Explicación: Según el tema introductorio al aprendizaje automático existen esos 3 tipos, los cuales se detallan levemente en la misma diapositiva. (Diapositiva 5)

1. En el aprendizaje bayesiano, el máximo a posteriori o MAP es:

{

= hMAP = argmaxP(h|D)

~ hMAP = argmaxP(D|h)

~ ninguna de las anteriores

}

Explicación: en el aprendizaje bayesiano se busca la hipótesis (h), de entre todas las H posibles, más probable si hemos observado una serie de datos D(máximo a posteriori o MAP). La segunda respuesta no equivale al máximo a posteriori o MAP, ya que sería correcta para P(h) = cte y máxima verosimilitud.

4. ¿Por cuántas fases pasa el aprendizaje bayesiano?:{

~ Por 3 fases.

= Por 2 fases.

~ El aprendizaje bayesiano no pasa por ningún tipo de fase.

}

Explicación: El aprendizaje bayesiano requiere de 2 fases: aprendizaje y clasificación. En aprendizaje se realizan cálculos previos y en clasificación hallamos la estimación MAP. En las transparencias 11 y 12 del tema 8 podemos observar esto.

5. El clasificador bayesiano "naive" asume que:{

~ los atributos que describen a los ejemplos son condicionalmente dependientes entre sí con respecto al concepto que se pretende aprender

=  los atributos que describen a los ejemplos son condicionalmente independientes entre sí con respecto al concepto que se pretende aprender

~ algunos de los atributos que describen a los ejemplos son condicionalmente dependientes entre si.

}

Explicación: El clasificador Bayesiana *naive*, también llamado a veces *idiot Bayes*, asume que los valores de los atributos son condicionalmente independientes dado el valor de la clase.

Osea: $P(a_1, \ldots, a_n \mid v_j) = \prod_{i} P(a_i \mid v_j)$

Por lo que:

\begin{displaymath}v_{NB} = argmax_{v_j \in V} \left( P(v_j) \prod_{i} P(a_i \mid v_j)  \right) \end{displaymath}

Los valores $P(a_i \mid v_j)$se estiman con la frecuencia de los datos observados.

6. Considerando 3 variables A=1, B=5, C=2, el resultado devuelto por argmax(A, B, C) es :{

~ 5

= B

~ C

}

Explicación: Al igual que max devuelve el máximo de los valores (en este caso sería 5), argmax devuelve la variable de la lista que maximiza el valor, en este caso B

7. Respecto a los fundamentos MAP y ML es cierto que  :{

~ MAP implica un máximo a priori.

= ML implica máxima verosimilitud, debido a que P(h) = cte.

~ Como P(D) es constante, depende de h.

}

Explicación: La primera opción y la tercera son incorrectas puesto que el MAP implica máximo a posteriori, el a priori es la característica del conocimiento al combinarlo con los datos de ejemplo, siendo esto una característica del aprendizaje bayesiano. (Diapositiva 2 y 4).

La tercera opción sería correcta si se indicase como independiente, tal como viene en la diapositiva 4.

La opción correcta, también de la diapositiva 4, nos viene explicada en el desglose de hMAP conforme se van marcando como constante P(D) y P(h).

8. Respecto al aprendizaje bayesiano, ¿cuál de las siguientes opciones es correcta?: {

~ Usado como clasificador, puede obtener la certeza de pertenecer a una clase.

~ Permite combinar los datos de ejemplo con conocimiento a posteriori.

= Permite combinar los datos de ejemplo con conocimiento a priori.

}

Explicación: Usado como clasificar, se obtiene la probabilidad (no la certeza, no tiene por qué pertenecer completamente) a cada clase del sistema (y no a una sola). Además, permite combinar los datos de ejemplo con conocimiento a priori, y no a posteriori.

9. En el Aprendizaje Supervisado  {

~ El conjunto de entrenamiento se usa para estimar el error de generalización

~ Hay que entrenar hasta alcanzar un error de entrenamiento mínimo

= Hay que entrenar hasta alcanzar el error de validación mínimo.

}

Explicación: La solución es la c) porque es la única que se ajusta a su función. La b), entrenar hasta alcanzar un error de entrenamiento mínimo, produciría sobreentrenamiento (overfitting). La a) no es porque el conjunto de entrenamiento se usa para ajustar los parámetros de aprendizaje del clasificador.

10. El conjunto de validación en un proceso de aprendizaje supervisado sirve

{

~ Para probar y depurar el proceso de aprendizaje.

= Para comprobar la capacidad de generalización del proceso y evitar el sobre entrenamiento.

~ Para realizar el entrenamiento en métodos de aprendizaje supervisado

}

Explicación. El conjunto de datos disponible para realizar un aprendizaje supervisado se divide en dos conjuntos: el de entrenamiento y el de validación. El primero sirve para realizar el proceso de aprendizaje en sí. El segundo son permite comprobar el funcionamiento del método tras el entrenamiento al enfrentarlo a datos que no han sido 'observados' durante el entrenamiento. Así comprobamos la capacidad de generalización y podemos detectar problemas de sobreentrenamiento.

11. En el aprendizaje supervisado, la regresión: {  
~ Calcula una función continua.  
= Estima una función continua.  
~ Ninguna de las anteriores.  
}  
Explicación: En la Transparencia 5 vemos es una caracteristica del aprendizaje supervisado.

12. Respecto a las hipótesis en el clasificador bayesiano:

{

~ El ejemplo el cual pertenece a la hipótesis está caracterizado como tuplas de atributos.

~ Las hipótesis son las clases a las que puede pertenecer un ejemplo

= Las dos respuestas son correctas

}

Explicación: Mediante la estimación de la probabilidad con la cual suponemos que los valores de los atributos son condicionalmente independientes para una clase dada observamos que el clasificador bayesiano tiene esas dos propiedades.

13. ¿Qué permite hacer el aprendizaje bayesiano?:{

~ permite combinar los datos de ejemplo con conocimientos a posteriori

~ ninguna respuesta es correcta

= permite combinar los datos de ejemplo con conocimientos a priori

}

Explicación: como podemos observar en la transparencia 2 del tema 8 usa conocimientos a priori para combinar estos datos

14. De los siguientes tipos de aprendizaje, indica la definición correcta: {

~Aprendizaje no supervisado: Tenemos una medida de lo bien o mal que está funcionando el algoritmo, pero no sabemos exactamente qué falla.

~Aprendizaje por refuerzo: Tenemos un conjunto de datos que queremos agrupar en clusters.

=Aprendizaje supervisado: Aprendemos a partir de ejemplos conocidos (etiquetados según su clase).

}

Explicación: En el Aprendizaje Supervisado aprendemos a partir de ejemplos conocidos (etiquetados según su clase).

Conjunto de validación: Aleatoriamente se parte el conjunto inicial de ejemplos en dos grupos.

Uno se usa como conjunto de entrenamiento, para ajustar los parámetros de aprendizaje del clasificador.

El otro es el conjunto de validación y se usa para estimar el error de generalización.

Se entrena hasta alcanzar el error de validación mínimo.

Entrenar hasta alcanzar un error de entrenamiento mínimo produciría sobreentrenamiento (overfitting).

Cuando los conjuntos de entrenamiento no son lo suficientemente grandes usamos la técnica de validación cruzada (k-fold cross validation).

15. El decisor de máxima verosimilitud, ML, asume que todas las hipótesis son equiprobables: {

= A priori.  
  
~ A posteriori.  
  
~ Durante el proceso de cálculo de máximo versimilitud.

}

Explicación: En ML (maximum likelihood) a priori, todas las hipótesis son igualmente probables.  
Transparencia 4.

16. El uso de reconocimiento de patrones a partir de ejemplos conocidos es una característica del:{

= Aprendizaje supervisado.  
~ Aprendizaje no supervisado.  
~ Aprendizaje por refuerzo.

}

Explicación: En el aprendizaje supervisado: conocemos la clase a la que pertenece cada ejemplo y aprendemos a partir de ejemplos conocidos (etiquetados según su clase).

17. ¿En que consiste el aprendizaje automatico? :{

~ Como una computadora no puede mejorar con la experienza, el aprendizaje automatico consiste en programar una computadora para que mejore  
en la realización de una tarea a partir de datos de ejemplo.

= Consiste en programar una computadora para que mejore  
en la realización de una tarea a partir de datos de ejemplo o  
de la experiencia

~ Una computadora programada solo es capaz de mejorar mediante la experiencia, los datos de ejemplo son ineficaces.

}

Explicación: Por definición directa es la segunda. Aqui nos dice que los datos de ejemplo son ineficaces, esto es absurdo, debido a que cuantos mas datos tenga desde un principio menor será el grado de aprendizaje que le costara funcionar de forma correcta. Esto me lleva a la respuesta 1 para rebocarla tambien debido a que nosotros podemos hacer que una computadora aprenda simplemente haciendo que almacene nuevos factores de cambio.

18. El aprendizaje por refuerzo, cual de las siguientes es correcta:{

=No parte de un conjunto de datos ejemplo.

~Parte de un conjunto de datos ejemplo (Como los supervisados).

~Se trata de un modelo mixto, donde existe un conjunto de ejemplos y además tenemos una forma para averiguar la calidad de un algoritmo.

}

Explicación: El aprendizaje por refuerzo se basa en una función que nos indica lo bien o mal que está funcionando el algoritmo. No nos basamos en información previa.

19. En la fase 1 de aprendizaje voc significa:{

= Conjunto de palabras en X (sin repetición y sin considerar preposiciones, artículos, etc.

~ El conjunto de posibles categorías.

~ Ninguna de las anteriores.

}

Explicación: Fase 1: Aprendizaje  
- Tomar un conjunto de ejemplos xi ∈ X etiquetados con las clases a las que pertenecen  
- Voc = conjunto de palabras en X (sin repetición y sin considerar preposiciones, artículos, etc.)

20. Dada esta definicion: "Es un método de aprendizaje Autómatico donde un modelos es ajustado

a las observaciones y no tiene ningún conocimiento a priori" a cual de las siguientes corresponde

{

~Aprendizaje supervisado

=Aprendizaje no supervisado

~Aprendizaje por refuerzo

}

Explicación: En el aprendizaje no supervisado no tenemos ningún dato de entrenamiento porque no tenemos ningún dato a priori y nos basamos en las observaciones para aprender.

21. El conjunto de validación:{

= Se emplea para estimar el error de generalización.

~ Se usa para ajustar los parámetros de aprendizaje del clasificador

~ Se trata de un conjunto que contiene los valores máximos

}

Explicación: Una vez estimada la generalización usaremos un conjunto de validación (conjunto aleatorio con muestras significativas) para observar el error.

22. En el aprendizaje supervisado ,aleatoriamente se parte el conjunto inicial de ejemplos en :{

~ dos grupo, series de pruebas y conjunto de validación.

= dos grupos ,conjunto de entrenamiento y conjunto de validación.

~  tres grupos ,conjunto de entrenamiento ,series de pruebas y conjunto de validación.

}

Explicación: Aleatoriamente se parte el conjunto inicial de ejemplos en dos grupos.Uno se usa como **conjunto de entrenamiento**, para ajustar los parámetros de aprendizaje del clasificador y el otro es el **conjunto de validación** y se usa para estimar el error de generalización.

23. He rectificado mi pregunta por estar ya planteada la anterior.

¿Cómo solucionamos el problema de sobreentrenar un algoritmo de aprendizaje? {

~ Mediante la validación cruzada.

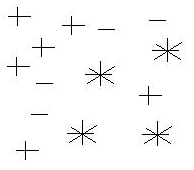
= A través del error de validación.

~ No tiene solución.

}

Explicación: El algoritmo de aprendizaje debe llegar a un estado en el que pueda saber el resultado de ciertos casos a partir de lo aprendido  con los datos de entrenamiento, generalizando para poder resolver situaciones distintas. Pero cuando un sistema entrena demasiado o con datos nuevos, el error de validación aumenta considerablemente y el éxito al responder empeora. Por tanto, se debe entrenar hasta que el error de validación sea mínimo y no se produzca un incremento mayor.

24. Dada la siguiente imagen y teniendo en cuenta que todos los atributos estan etiquetados y clasificados, podemos decir que:



{

= Tiene 3 tipos clases, una dimensión por cada atributo y pertenece al tipo de aprendizaje Supervisado.

~ Tiene 1 tipo de clase, 3 dimensiones por cada atributo y pertenece al tipo de aprendizaje No Supervisado.

~ No tiene ninguna clase, ninguna dimensión por cada atributo y pertenece al tipo de aprendizaje Supervisado.

}

Explicación: Tenemos 3 tipos de clases: +, -, \* ,una dimensión por cada atributo,  y pertenece al grupo de aprendizaje supervisado porque está clasificado y etiquetados.

25. Un ejemplo de área de aplicación del aprendizaje automático es: {

= Minería de datos (Data mining)

~ Máquinas de estados (State machines).

~ Árboles de comportamiento (Behavior trees).

}

Explicación: Un ejemplo válido de aplicación de aprendizaje automático es la minería de datos ya que trata de obtener conclusiones a partir del análisis estadístico de grandes cantidades de datos, por lo tanto, el uso del aprendizaje automático es muy favorable en este campo de la informática.

26. Ante un problema, aplicaremos el clasificador bayesiano naive cuando: {

~ Dispongamos de conjuntos de entrenamiento de tamaño medio o grande.

~ Los atributos que describen a los ejemplos sean independientes entre sí con respecto al concepto que se pretende aprender.

= Ambas son correctas.

}

Explicación: En la transparencia 6 del tema 8, podemos observar que cuando tenemos un conjunto de entrenamiento pequeño la estimación al emplear naive bayes es inadecuada. Y suponemos que los valores de los atributos son condicionalmente independientes para una clase dada.

27. ¿Cuál de los siguientes conjuntos del aprendizaje supervisado no es el correcto?  :{

~ Entrenamiento

= Cluster

~ Validación

}

Explicación: Los conjuntos del aprendizaje supervisados son el 'entrenamiento' y 'validación' por lo tanto la opción correcta sería 'cluster' ya que esta forma parte del aprendizaje no supervisado.

El conjunto de entrenamiento ajusta los parámetros de aprendizaje del clasificador. (Aprendizaje supervisado)  
El conjunto de validación se usa para estimar el error de generalización. (Aprendizaje supervisado)  
La agrupación cluster forma parte de un conjunto de datos. (Aprendizaje no supervisado)

Para más detalle: Pág 5 y 6 del pdf "Introducción al aprendizaje automático"

28. Cuando los conjuntos de entrenamiento no son suficientemente grandes utilizamos la técnica de validación cruzada k-fold cross validation, que funciona: {

~ Coge un conjunto que contiene k elementos y lo divide en 2 partes del mismo tamaño, uno para validación y otro para entrenamiento.

~ Coge k conjuntos de elementos sin importar el tamaño y coje uno de esos conjuntos de forma aleatoria para que valide, y los otros restantes (k-1) serán para entrenamiento.

= Coge un conjunto y lo divide en k conjuntos del mismo tamaño y hace un conjunto para validación y los otros para entrenamiento, cambiándose los papeles para que todos sean validadores y de entrenamiento.}

Explicación: Tenemos un numero pequeño y tenemos que quitar unos pocos para entrenar aún se queda más pequeño.  
Lo divido en varios conjuntos iguales de pequeña cantidad, entonces coges un conjunto de elementos que lo utilizas para validar y el resto para entrenar,  
después haces lo mismo con el siguiente grupo, y así con todos los ejemplos, de forma que cuando tienes pocos no quitas ejemplos para validar.

29. En el clasificador bayesiano:

{

~ Las hipótesis son las clases a las que puede pertenecer un ejemplo.

~ Suponemos ejemplos caracterizados como tuplas de atributos <a1, a2, ... an>

= Las dos son correctas.

}

Explicación: Las dos son correctas tal y como podemos comprobar en la página 5 del tema 8.

30. Cuál de las siguientes afirmaciones es correcta: :{

~ En el aprendizaje supervisado, en caso de no haber un gran número de ejemplos, se aplica la técnica k-fold cross validation

~ En el aprendizaje supervisado, se divide el conjunto inicial de ejemplos en dos grupos , uno para entrenar y otro para la validación

=  Ambas son correctas

}

Explicación: El aprendizaje supervisado se basa en el entrenamiento y posterior validación para estimar el error de generalización. Por otra parte, pueden existir situaciones en las que no podamos obtener suficientes ejemplos para ambos conjuntos; en ese caso dividiremos los ejemplos en k conjuntos, probando k-1 como entrenamiento y 1 como validación, probando todas las combinaciones posibles.

31. Indica cuál de las siguientes afirmaciones sobre el teorema de bayes es correcta: {

~ El aprendizaje bayesiano se divide en 5 partes

= Es utilizado para detectar si un correo es spam o no mediante palabras guardadas en un vocabulario

~ Ninguna de las siguientes afirmaciones es correcta

}

Explicación: El teorema de Bayes es uno de los métodos más utilizados para detectar si nuestro correo es spam; ya que comprueba si existen palabras que tenemos guardado en un vocabulario, pudiendo así decidir si el correo es spam o no lo es.

32. Respecto al clasificador bayesiando podemos decir que es correcto que:{

~ Las hipótesis son las fases del aprendizaje.

= Las hipótesis son las clases a las que puede pertenecer un  
ejemplo.

~ ninguna de las anteriores

}

Explicación: Como podemos ver en la página 5 del tema 8, las hipótesis son las clases a las que puede pertenecer un ejemplo. Suponemos ejemplos caracterizados como tuplas de atributos <a1, a2, ... an>.

33. Cuando los conjuntos de entrenamiento no son lo suficientemente grandes usamos la técnica de...: {

~ Validación paralela.

~ Conjunto de validación.

= Ninguna de las anteriores es correcta.

}

Explicación: Cuando los conjuntos de entrenamiento no son lo suficientemente grandes usamos la técnica de validación cruzada (k-fold cross validation).

34. Dentro del tipo de aprendizajes, cual es el que agrupa los datos en clusters {

= Aprendizaje no supervisado

~ Aprendizaje supervisado

~ Aprendizaje por refuerzo

}

Explicación: El aprendizaje supervisado asigna los datos según su categoría y ya conoce previamente la clasificación, mientras que el aprendizaje por refuerzo sabe que algo falla pero no sabe clasificarlo porque lo desconoce.

35. Cuando estamos en la fase de aprendizaje en un clasificador de textos usando naive bayes, por qué aparece el 1 en el numerador y el término |Voc| en el denominador del cálculo de la probabilidad de una palabra en una categoría dada: P(wk|cj) = (nk+1)/(n+|Voc|) :{

~ Para evitar que la probabilidad salga 1 si la palabra no ha aparecido nunca.

= Para evitar que la probabilidad salga 0 si la palabra no ha aparecido nunca.

~ Para que el cálculo de la probabilidad sea lo más ajustado posible a la realidad.}

Explicación: Si una palabra no ha aparecido nunca en cj la probabilidad nos daría 0 con lo que nunca diríamos que el texto pertenece a cj si aparece la palabra.

36. Estamos implementando un modelo de aprendizaje para guiar a nuestro robot autómata "Emilio" en un entorno laberintico mediante sucesivas pruebas a base de prueba/error; y utlizando simplemente 3 reglas de movimiento, las cuales impiden retroceder en el mapa, y que son: izquierda, adelante y derecha. Únicamente  podemos avanzar, de modo que no podemos ir hacia atrás en el mapa, ni usando una regla específica (como se ha comentado), ni usando giros a la izquierda o derecha. Sabiendo esto, indica que esquema de aprendizaje se adaptaría más al modelo planteado:

{

~ Aprendizaje supervisado

~ Aprendizaje NO supervisado

= Ninguna de las anteriores

}

Explicación: Correspondería al modelo de aprendizaje por refuerzo. Ya que para evitar caminos que no lleven a la salida en el laberinto, partimos de un entorno desconocido del cual no poseemos conocimiento previo; y por tanto necesitamos realizar sucesivas pruebas para determinar la solución a base de prueba/error. Además descartamos el aprendizaje supervisado y no supervisado; por ser modelos que parten de ejemplos conocidos y de recolección de datos (data mining) respectivamente. Los cuales no se adaptan al problema planteado.

37. El aprendizaje de bayesiano usado como clasificador:{

~ puede obtener probabilidades de no pertecener a ninguna clase

= puede obtener probabilidades de pertenecer a cada clase.

~ ninguna de las anteriores

}

Explicación: la respuesta correcta es la segunda y se puede ver en la pagina 2 del tema.

38. ¿Qué necesidades impulsan la idea del aprendizaje automático?:{

= Ambas son correctas.

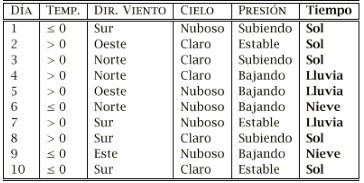
~ Situaciones o tareas donde el algoritmo que se requiere debe adaptarse a circunstancias particulares.

~ Algoritmos difíciles de programar "a mano".

}

Explicación: Por separado, segunda y tercera son correctas pero la "más correcta" sería la primera ya que engloba a ambas. Como vimos en el tema algoritmos que deban adaptarse existen varios tales como la minería de datos, detección de Spam, reconocimiento del habla, etc. este último también relacionado con aquellos dificiles de programar, al igual que el reconocimiento de escritura manuscrita, de objetos...

39. Sea el ejemplo de la predicción del tiempo visto en clase:



Podemos decir que la predicción del tiempo para

<presión=Subiendo,cielo=Nuboso>

sería: {

~ Lluvia

~ Nieve

= Sol

}

explicación:

Por lo tanto:

Sol:

P(Sol) = 0,5

P(presion=Subiendo|Sol) = 3/5

P(cielo=Nuboso|Sol) = 1/5

**0,5\*(3/5)\*(1/5) = 3/50**

Lluvia:

P(Lluvia) = 0,3

P(presion=Subiendo|Lluvia) = 0

P(cielo=Nuboso|Lluvia) = 2/3

**0,3\*0\*(2/3) = 0**

Nieve:

P(Nieve) = 0,2

P(presion=Subiendo|Nieve) = 0

P(cielo=Nuboso|Nieve) = 2/2 = 1

**0,2\*0\*1 = 0**

**Como 3/50 es mayor que 0, podemos predecir que, para <presión=Subiendo,cielo=Nuboso>, hará SOL.**

40. Un alumno de la UA guarda los resultados de sus notas finales en 1ª convocatoria de las asignaturas que ha cursado, así como si la asignatura era de primer o segundo cuatrimestre, si el profesor que le impartia la teoria era titular o asociado y la base de conocimiento de dicha asignatura (Programación, sistemas o teoria de la información) en la siguiente tabla:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Asignatura | Base de Conocimiento | Cuatrimestre | Profesor | Nota final |
| A | Sistemas | Primero | Titular | Aprobado |
| B | Programación | Primero | Titular | Sobresaliente |
| C | Sistemas | Primero | Titular | Notable |
| D | Teoria de la información | Primero | Asociado | Suspenso |
| E | Programación | Primero | Titular | Sobresaliente |
| F | Sistemas | segundo | Titular | Aprobado |
| G | Teoria de la información | segundo | Asociado | Notable |
| H | Programación | segundo | Asociado | Sobresaliente |
| I | Programación | segundo | Titular | Suspenso |
| J | Sistemas | Primero | Titular | Notable |
| K | Teoria de la información | Primero | Titular | Aprobado |
| L | Programación | Primero | Asociado | Sobresaliente |

Si para este curso se ha cogido una asignatura de programación de segundo cuatrimestre y la teoria se la imparte un profesor asociado ¿Qué calificación es más probable que obtenga en esta asignatura?

{

= Sobresaliente.

~ Aprobado.

~ Suspenso.

}

Explicación: Que mejor explicación que los cálculos realizados.

P(Sobresaliente) = 4/12 = 1/3

P(BC = Programación | Sobresaliente) = 1

P(P = Asociado | Sobresaliente) = 1/2

P(C = Segundo | Sobresaliente) = 1/4

P(Programación, Asociado, Segundo | Sobresaliente) = 1/3 \* 1 \* 1/2 \* 1/4 = 1/24

P(Notable) = 3/12 = 1/4

P(BC = Programación | Notable) = 0

P(Programación, Asociado, Segundo | Notable) = 0

P(Aprobado) = 3/12 = 1/4

P(BC = Programación | Aprobado) = 0

P(Programación, Asociado, Segundo | Aprobado) = 0

P(Suspenso) = 2/12 = 1/6

P(BC = Programación | Suspenso) = 1/2

P(P = Asociado | Suspenso) = 1/2

P(C = Segundo | Suspenso) = 1/2

P(Programación, Asociado, Segundo | Suspenso) = 1/6 \* 1/2 \* 1/2 \* 1/2 = 1/48

41. ¿Qué tipo de aprendizaje funciona a partir de ejemplos conocidos? :{

= Aprendizaje supervisado

~ Aprendizaje no supervisado

~ Aprendizaje por refuerzo

}

Explicación:

El aprendizaje supervisado se caracteriza porque conocemos la clase a la que pertenece cada ejemplo (asignamos una clase de un conjunto de clases discreto).

El aprendizaje no supervisado agrupa los datos en clusters

El aprendizaje por refuerzo, establece lo bien o mal que funciona el algoritmo

T8, página 5 y 6

42. En el aprendizaje bayesiano durante la fase 1 (aprendizaje), después de tomar un conjunto de ejemplos xi ∈ X etiquetados con las clases a las que pertenecen, debemos calcular la probabilidad a priori P(cj) para cada clase cj, lo cual se realiza de la siguiente manera:

{

~ P(cj) = nº de ejemplos no etiquetados con cj / nº de ejemplos etiquetados con cj

= P(cj) = nº de ejemplos etiquetados con cj / nº total de ejemplos

~ P(cj) = nº total de ejemplos / nº de ejemplos etiquetados con cj

}

Explicación: visto así puede parecer una pregunta compleja, pero únicamente estoy preguntando como calcularíamos, por ejemplo en el ejercicio de predicción del tiempo, la probabilidad de la clase cj = Sol.

A partir del conjunto de ejemplos que tenemos, calculamos el total de éstos que pertenecen a esa clase y lo dividimos por el nº total de ejemplos |X| (como podemos ver en la diapositiva 8 del tema 8 o en diversos ejercicios propuestos por los compañeros)

La pregunta hace referencia a la página 11, a la fórmula P(cj) = |docs j| / |X| que está aplicada al ejemplo del clasificador de textos.

43. En el aprendizaje supervisado:  {

~ tenemos una medida de cómo está funcionando el algoritmo, sin saber exactamente qué falla.

= conocemos la clase a la que pertenece cada ejemplo.

~ tenemos un conjunto de datos que agrupamos en clusters.

}

Explicación: Con el aprendizaje supervisado aprendemos a partir de ejemplos conocidos etiquetados según su clase.

44. La **minería de datos** intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automatico,estadistica y sistemas de BD . Consiste en extraer información de un conjunto de datos y transformarla en una estructura comprensible para su uso posterior.

Podemos decir que NO utiliza el tipo de aprendizaje:{

= Refuerzo.

~ Supervisado.

~No utiliza ninguno de los anteriores.

}

Explicación: La mineria de datos es un ejemplo clasico que utiliza el tipo de aprendizaje Supervisado, es una técnica para deducir una función a partir de datos de entrenamiento.No utiliza el aprendizaje por refuerzo.

45. Dados los tipos de aprendizaje y en concreto el Aprendizaje por Refuerzo, señala la afirmación correcta:{

~ En el Aprendizaje por Refuerzo tenemos una medida de lo bien que está funcionando el algoritmo y sabemos qué falla exactamente.

~ En el Aprendizaje por Refuerzo tenemos una medida de incertidumbre que nos da el algoritmo y sabemos aquello qué falla exactamente.

= En el Aprendizaje por Refuerzo tenemos una medida de lo bien o mal que está funcionando el algoritmo, pero no sabemos exactamente qué falla.

}

Explicación: Véase diapositiva nº5 de "Introducción al aprendizaje automático".

 46.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Día** | **tráfico** | **pasajeros** | **tiempo** | **llegada del autobus** |
| 1 | poco | muchos | lluvia | **puntual** |
| 2 | normal | pocos | sol | **puntual** |
| 3 | congestión | muchos | sol | **retrasado** |
| 4 | normal | pocos | nieve | **retrasado** |
| 5 | normal | muchos | lluvia | **retrasado** |
| 6 | poco | pocos | sol | **temprano** |

Cuáles son las predicciónes correctas de la llegada del autobus para <pasajeros = muchos, tiempo = sol> utilizando *P( ci ) ∏j P ( aj ∣ ci )*?

{

~ *ci* = puntual → 1/9, *ci* = temprano → 0, *ci* = retrasado → 1/18

= *ci* = puntual → 1/12, *ci* = temprano → 0, *ci* = retrasado → 1/9

~ *ci* = puntual → 1/12, *ci* = temprano → 1, *ci* = retrasado → 1/9

}

Explicación:

P( ci ) ∏j P ( aj ∣ci ) à P(ci) P(pasajeros = muchos | ci ) P (tiempo = sol)

* **ci = puntual:**

P(puntual) = 2/6 = 1/3

P(pasajeros = muchos | puntual) = 1/2

P(tiempo = sol | puntual ) = 1/2

P( ci ) ∏j P ( aj ∣ci ) = 1/3 \* 1/2 \* 1/2 = 1/12

* **ci = temprano:**

P(temprano) = 1/6

P(pasajeros = muchos | temprano) = 0

P(tiempo = sol | temprano ) = 1

P( ci ) ∏j P ( aj ∣ci ) = 1/6 \* 0 \* 1 = 0

* **ci = retrasado:**

P(retrasado) = 3/6 = 1/2

P(pasajeros = muchos | retrasado) = 2/3

P(tiempo = sol | retrasado ) = 1/3

P( ci ) ∏j P ( aj ∣ci ) = 1/2 \* 2/3 \* 1/3 = 1/9

47. En el Aprendizaje por refuerzo  {

~ conocemos la clase a la que pertenece cada ejemplo.

= tenemos una medida de lo bien o mal que está funcionando el algoritmo, pero no sabemos exactamente qué falla.

~ tenemos un conjunto de datos que queremos agrupar en clusters.

}

Explicación: Como podemos ver en la transparecia 6 del documento "Introducción al aprendizaje automático" la solución correcta es la segunda, ya que la primera corresponde al aprendizaje supervisado y la tercera al no supervisado.

48. acerca de las características de las redes de aprendizaje bayesianas, es dicho:  
{  
~Un conocimiento a priori puede ser combinado con los datos observados para determinar la probabilidad de una hipótesis.  
~Los métodos bayesianos pueden acomodar las hipótesis que hacen predicciones probabilísticas.  
=Ambos son correctas.

}

Explicación: Las decisiones óptimas se pueden tomar en estas probabilidades, junto con los datos observados.

49. Dadas las variables X=9, Y=3, Z=5, la solución a max(X,Y,Z) y a argmax(X,Y,Z) es :{

= 9 y X respectivamente.

~ X y 9 respectivamente.

~ Ambas son correctas puesto que al fin y al cabo se refieren a lo mismo.

}

Explicación: La función max devuelve el mayor valor de las 3 variables y la función argmax devuelve la variable con mayor valor.

50. Un algoritmo de aprendizaje: {

= es consistente si obtiene una hipótesis que no comete ningún error sobre los ejemplos de entrenamiento.

~ es inconsistente si obtiene una hipótesis que no comete ningún error sobre los ejemplos de entrenamiento.

~ es consistente si obtiene una hipótesis que comete algún error sobre los ejemplos de entrenamiento.

}

Explicación: Por definición, Un algoritmo de aprendizaje es consistente si obtiene una hipótesis que no comete ningún error sobre los ejemplos de entrenamiento.

51. En el tipo de aprendizaje no supervisado tenemos: {

~ una medida de lo bien o mal que está funcionando el algoritmo.

= un conjunto de datos que queremos agrupar en clusters.

~ El aprendizaje no supervisado no pertenece a los tipos de Aprendizaje.

}

*Explicación: En la diapositiva 5 del tema de Introducción al aprendizaje. Se definen tres tipos de aprendizaje, en concreto el aprendizaje no supervisado se define de la siguiente forma se tiene un conjunto de datos que queremos agrupar en clusters.*

52. Después de realizar el cálculo del siguiente caso de predicción meteorológica, obtenemos P(T<0,Viento=Norte|Lluvia)=0 ¿Qué podemos determinar? {  
~ No puede ser 0, no es una probabilidad.  
~ Es seguro que llueva.  
= No lloverá.  
}  
  
Explicación: Con en listado actual de casos no hemos podido determinar la probabilidad de alguno de los casos de la predicción. Por tanto, al no haber un caso previo, determinaremos que no lloverá si la temperatura es menor de 0º y la dirección del viento viene del norte.

53. Cuál de las siguientes afirmaciones acerda del aprendizaje bayesiano es falsa:

{

~ Permite combinar los datos de ejemplo con conocimiento a priori

~ Usado como clasificador, puede obtener probabilidades de pertenencer a cada clase

= Esta basado en el teorema --> P(h|D) = P(D) / P(D|h)P(h)

}

*Explicación: La falsa es la ultima puesto que el teorema de bayes esta mal expresado, tendria que ser --> P(h|D) = P(D|h)P(h) / P(D)*

54. Google completa la cadena que introduces en la barra del buscador, utiliza el modelo oculto de markov que es un modelo estadistico cuyo objetivo es determinar los parámetros desconocidos de dicha cadena a partir de los parámetros observables,  permite combinar los datos de la cadena con conocimiento a priori.



Podemos decir que :{

~ No se basa en ningun tipo de aprendizaje automático.

= Se basa en el tipo de aprendizaje supervisado.

~ Se basa en el tipo de aprendizaje no supervisado.

}

Explicación: es supervisado porque como dice en el enunciado permite combinar los datos de la cadena con conocimiento a priori (caracteristica de el aprendizaje bayesiano (supervisado))

55. Cuál de las siguientes afirmaciones es correcta, sobre los aprendizajes automáticos:{

= En el aprendizaje supervisado el error de generalización se estima mediante un conjunto de validación.

~ En el aprendizaje por refuerzo se utilizan clusters para agrupar un conjunto de datos.

~ En el aprendizaje supervisado se entrena hasta alcanzar un error de entrenamiento mínimo.

}

Explicación:  El primero es correcto, el conjunto de validación se utiliza para estimar el error de generalización. Los clusters se utilizan en el aprendizaje no supervisado y por último, en el aprendizaje supervisado, se entrena hasta alcanzar un error de validación.

56. Respecto al aprendizaje no supervisado, donde tenemos un conjunto de datos que queremos agrupar en clusters: {

~ El conjunto de datos a clasificar está etiquetado.

= No tenemos etiquetas de a que pertenecen los datos.

~ buscamos una acción a realizar y no una clasificación de ejemplos.

}

*Explicación:*

*La opción correcta es que los datos a clasificar no están etiquetados (segunda afirmación), como por ejemplo en la minería de datos (data mining) donde queremos extraer conocimiento de un conjunto donde a simple vista parece no haber relación. La primera respuesta es precisamente el punto de partida de un aprendizaje supervisado (todos los datos pertenecen a un conjunto de clases discreto), mientras que la tercera respuesta está relacionada con el aprendizaje por refuerzo.*

57. En una competición de tiro con arco después de 10 lanzamientos de un competidor se han registrado los siguientes datos:

|  |  |  |  |
| --- | --- | --- | --- |
| Tiro | Viento | Distancia (m) | Puntuación |
| 1 | SUR | 50 | 30 |
| 2 | ESTE | 60 | 10 |
| 3 | SUR | 50 | 20 |
| 4 | NORTE | 60 | 30 |
| 5 | NORTE | 50 | 10 |
| 6 | ESTE | 75 | 30 |
| 7 | SUR | 60 | 30 |
| 8 | SUR | 60 | 10 |
| 9 | OESTE | 50 | 30 |
| 10 | ESTE | 60 | 20 |

Calcula la predicción de la puntuación para <viento = SUR, distancia = 60> {

~ c(30) = 0.04, c(20) = 0.05, c(10) = 0.0333

~ c(30) = 0.08, c(20) = 0.05, c(10) = 0.0333

= c(30) = 0.08, c(20) = 0.05, c(10) = 0.0666

}

Explicación:

Usando la formula: P(c) P(viento = SUR | c) P(distancia = 60 | c)

c(30) → 5/10 \* 2/5 \* 2/5 = 0.08

c(20) → 2/10 \* 1/2 \* 1/2 = 0.05

c(10) → 3/10 \* 1/3 \* 2/3 = 0.0666

58. Pregunta: Según el aprendizaje bayesiano ¿Cómo se estima una probabilidad? {

~ nº total de ejemplos / nº de ejemplos de la clase actual

~ nº de ejemplos con atributos a1, a2, ..., an / nº de ejemplos de la clase actual

= nº de ejemplos de la clase actual / nº total de ejemplos

}

Solución: Para estimar una probabilidad debemos dividir el número de ejemplos de la clase entre el número total de ejemplos.

59. Acerca de las dificultades del aprendizaje bayesiano{

 ~ Necesidad de un conocimiento a priori. Si no se tiene este conocimiento estas  probabilidades han de ser estimadas

~Coste computacional alto. En el caso general es lineal con el número de hipótesis candidatas.

= Ambas son ciertas

}

Explicación: Como ya sabemos en el aprendizaje bayesiano, hay que conocer la probabilidad de los hipótesis a priori con el fin de eliminir facilmente las hipótesis incompatibles. El coste computacional puede ser muy elevido debido al número de hipótesis candidatas.

60. ¿Cuál es la idea del aprendizaje automático?{  
= En programar una computadora para que mejore en la realización de una tarea a partir de datos de ejemplo o de la experiencia;

~ Hacer uso delconocimiento especializado para resolver problemas como  
un especialista humano.

~ De cuanto ayuda elconocer el valor de una variable aleatoria X para  
conocer el verdadero valorde otra Y.

}  
Explicación:  
La opción correcta es la primera, siendo que la segunda es el concepto del sistemas expertos e la ultima es el concepto del ganancia de información.

61. En "naive bayes" el cálculo simplificado consiste en: {

~ Calcular la probabilidad de que se den los atributos del ejemplo a clasificar condicionada a la clase que se esta probando.  
~ Calcular el sumatorio de probabilidades de todos los atributos del ejemplo a clasificar condicionado a la clase que se esta probando.

= Calcular el productorio de probabilidades de todos los atributos del ejemplo a clasificar condicionado a la clase que se esta probando.

}  
Explicación: En "naive bayes" se considera que los atributos son condicionalmente independientes a una clase dada, por lo que en vez de calcular la probabilidad de que se de cierta combinación de atributos dentro de una clase, se calcula el productorio de la probabilidad de que aparezca cada uno de los atributos del ejemplo a clasificar en la clase dada.

62. ¿Cuál de estas opciones es correcta con respecto a la simplificación?:{

~ Se emplea para estimar el error de generalización.

= Suponemos que los valores de los atributos son  
condicionalmente independientes para una clase dada

~ Ambas son correctas

}

Explicación: Diapositiva 6 del tema 8.

63. ¿Cuál de estas opciones es característica de el Aprendizaje Bayesiano?:{

~ No es necesario conocimiento a piori.

= Es necesario conocimiento a piori.

~ Tiene un bajo costo computacional.

}

Explicación: Página 2 del sitio web <http://gaiatemp.fdi.ucm.es/people/pedro/aa02/05bayesiano.pdf>.

64. ¿Cuál de las siguientes afirmaciones no corresponde al aprendizaje bayesiano?:

{

= Permite combinar los datos de ejemplo con conocimiento a posteriori

~Usado como clasificador, puede obtener probabilidades de pertenecer a cada clase

~Posibilidad de construir representaciones más complejas.

}

Solución: El aprendizaje bayesiano permite combinar los datos de ejemplo con conocimiento a priori.

65. ¿En cuál de las aprendizajes se considera que el problema de la construcción de hipótesis a partir de datos es un subproblema del problema más importante relacionado con la formulación de predicciones?:{

= Aprendizaje bayesiano

~ Aprendizaje de funciones linealmente separables

~ Aprendizaje por propagación posterior

}

*Explicación: Esta es la definición del aprendizaje bayesiano.*

66. Sabemos que el teorema de Bayes aplicado al aprendizaje permite conocer el máximo a posteriori MAP, por lo tanto, podemos señalar como opción incorrecta:{

~ El aprendizaje bayesiano consiste en buscar la hipotesis h ( de entre todas las H posibles) más probable si hemos observado una serie de datos D (MAP).

~ El aprendizaje bayesiano permite combinar los datos de ejemplo con conocimiento a priori.

= El aprendizaje bayesiano permite combinar los datos de ejemplo con conocimiento a posteriori.

}

Explicación: las 2 primeras opciones son correctas pero la última para ser correcta tendría que ser como la primera, con conocimiento a priori.

67. El parendizaje bayesiano se basa en el teorema de Bayes, el cual es:{

= P(h|D) = P(D|h) \* P(h) / P(D).

~ P(h) = P(D|h) \* P(h).

~ P(D) = P(h)\* P(D) / P(D|h).

}

*Explicación: El teorma de bayes es P(h|D) = P(D|h) \* P(h) / P(D), donde:*

*P(h) son las probabilidades a priori.*

*P(D|h) es la probabilidad de D en la hipótesis h.*

*P(h|D) son las probabilidades a posteriori.*

68. Respecto al aprendizaje bayesiano indica la opción correcta:{

~ En la fase 3 hallamos la estimación MAP.

~ En la fase 2 calculamos docsj , documentos de la clase cj.

= Ninguna de las anteriores.

}

Explicación: No existe fase 3. El aprendizaje bayesiano consta de 2 fases: aprendizaje y clasificación. El calculo de docsj se da en la fase 1.

69. En que consisten las simplificaciones.  {

~ No importa si aparecen todos los elementos de nuestro estudio, sino que importa el orden en el cual aparecen.

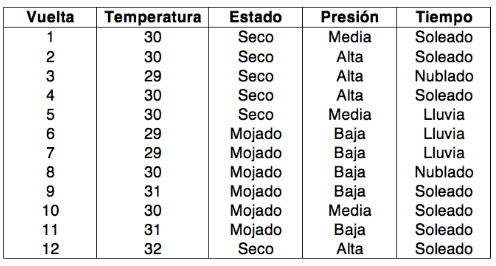
~ Los elementos a buscar son dependientes entre sí.

= No importa el orden de aparición de los elementos, sólo importa si estos elementos estan presentes.

}

Explicación: Por definición, la simplificación consiste en que el orden no importa, siendo lo realmente importante que el elemento se encuentre presente, tal y cómo indica en la página 10 del tema, aparte de que los elementos son independientes, tal y cómo aparece en la transparencia 6.

70. Los datos de un neumático de un Fórmula 1 en una carrera son los siguientes:



La predicción de una **alta** **presión** del neumático cuando el estado es **mojado** y el tiempo **soleado** es {

~ 0.58

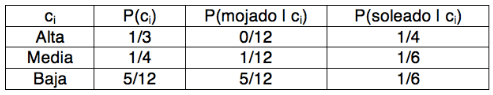
= 0

~ 0.25

}

Explicación:

La predicción correcta es 0, ya que nunca se va a dar una alta presión cuando el neumático esté mojado.



P(ci)P(mojado | ci)P(soleado | ci) = 1/3 \* 0 \* 1/4 = 0

71.  En el aprendizaje bayesiano, al estimar las probabilidades.  {

~ Suponemos que los valores de los atributos son condicionalmente dependientes para una clase dada.

~ Suponemos que los valores de los atributos son incondicionalmente dependientes para una clase dada.

= Suponemos que los valores de los atributos son condicionalmente independientes para una clase dada.

}

Explicación: Por definición, al estimar las probabilidades suponemos que los valores de los atributos son condicionalmente independientes para una clase dada.

Tema 8 - Pág. 6

72. Si hablamos del aprendizaje bayesiano, en la fase de clasificación: {

= Dado un documento x=w1,w2,...,wn, nos quedamos con las posiciones de palabras que están contenidas en Voc y devolvemos la estimación map.

~ Dado un documento x=w1,w2,...,wn, nos quedamos con las posiciones de palabras que no están contenidas en Voc y devolvemos la estimación map

~ Ambas son incorrectas. }

Explicacion: Transparencia 12 del tema8.

73. ¿Cuál de los siguientes afirmaciónes acerca de problemas de la aprendizaje bayesiano NO es correcta?

{

~ Necesidad de un conocimiento a priori.

~ Costo computacional alto.

= No hay la posibilidad de construir representaciones más complejas.

}

Explicación:

1. Si no se tiene conocimientos a priori estas probabilidades han de ser estimadas.

2. El costo computacional del aprendizaje de una red bayesiana desde datos, crece con el número de variables en la misma y con el número de casos, por consiguiente surge el problema de identificar algoritmos eficientes que aprendan desde los datos.

3.Exactamente el contrario es verdad y por eso no es un problema de la aprendizaje bayesiano

74. En el aprendizaje Bayesiano existe Maxima Verosimilitud (ML) cuando  {

~ P(h) = no cte

= P(h) = cte

~ No depende de P(h)

}

Explicación: Si suponemos que todas las hipotesis son igual de probables, P(h)=cte, entonces existe maxima verosimilitud

75. ¿Qué presupone el classificador bayesiano „naive“?{

~ Que todos los atributos tienen que ser dependientes uno del otro

~ Que todos los atributos siempre tienen dos nodos padres

= Que todos los atributos tienen que ser independientes

}

Explicación:

La respuesta correcta es la tercera. Los atributos de un clasificador bayesiano deben ser independientes entre sí. Solo dependen del conjunto de las posibles categorías.

76. La siguiente fórmula es la base del aprendizaje bayesiano:

h\_MAP = argmaxP(h | D)

¿Qué significa?:{

~ Buscar la hipótesis h con el valor máximo a posteriori.

~ Buscar la hipótesis h más probable si hemos observado una serie de datos D.

= Las tres respuestas son correctas.

}

Explicación: El aprendizaje bayesiano se basa en buscar la hipótesis h (de entre todas las H posibles) más probable si hemos observado una serie de datos D (máximo a posteriori o MAP), y se representa con la fórmula:

h\_MAP = argmaxP(h | D)

77. Podemos calcular la probabilidad de los atributos condicionada a las diferentes categorías  con la aproximación  considerando: {

= Los atributos son independientes entre si.  
  
~ La probabilidad de las clases se suponen las mismas.  
  
~ Ninguna de las anteriores.

}

Explicación: Solo si las probabilidades de atributos son independientes entre sí podemos transformar las probabilidades condicionales en productos de probabilidades.

78. En la fase 1, Aprendizaje, de Clasificador bayesiano “naive”, porque se añade el 1 a nk**(nk+1)**

{

~Porque las clases se relacionan con attributos independientes.

=Para que cuando la probabilidad dé 0 normalmente, ese valor 0, no influe demasiado en el resultado final.

~Niguna de las dos es cierta.

}

Explicación: Como bien lo explicó la profesora, se añade un 1 para que en caso de que tengamos como probabilidad real un 0, este valor no influe el resultado.

79. ¿Qué método de aprendizaje utilizaríamos para hacer que un robot con un brazo mecánico pudiese mejorar jugando a ping pong?:{  
  
= Aprendizaje por refuerzo  
  
~ Aprendizaje supervisado  
  
~ No se puede implementar ese tipo de aprendizaje en un sistema como el especificado.  
  
}  
  
Explicación: Según el ejemplo planteado en clase en la explicación de los tipos de aprendizaje, el aprendizaje por refuerzo es aquel que mide cómo está funcionando el algoritmo, y puesto que el juego se mejora con la experiencia y la técnica adquirida, el algoritmo se refina según avanza el juego con este tipo de aprendizaje.

80. El aprendizaje bayesiano: {

      =Basado en el teorema de Bayes

     ~ Permite combinar los datos de ejemplo con conocimiento a posteriori

     ~ Es incapaz de obtener probabilidades de pertenecer a cada clase.

}

     Explicación: El aprendizaje bayesiano está basado en el teorema de Bayes y permite combinar los datos de ejemplo cono conocimiento a PRIORI, además si se usa como clasificador si puede obtener las probabilidades de pertenecer a cada clase.

81. El aprendizaje bayesiano consiste en:{

~Buscar la hipótesis h más probable si hemos observado una serie de datos D. hMAP≡ argminP(h|D)

=Buscar la hipótesis h más probable si hemos observado una serie de datos D. hMAP≡ argmaxP(h|D)

~Buscar la hipótesis D más probable si hemos observado una serie de datos h. DMAP≡ argmaxP(h|D)

}

Explicación: Como estudiamos en la página 3 de las transparencias del tema8, el aprendizaje bayesiano consiste en buscar la hipótesis h (de entre todas las H posibles) más probable si hemos observado una serie de datos D (máximo a posteriori o MAP).  hMAP≡ argmaxP(h|D)

82. Es falso que un clasificador Bayesiano:{

~  Proporciona una clase etiquetada como salida a partir de un conjunto de características tomadas como entradas.

~ Calcula la probabilidad de cada hipótesis de los datos y realiza predicciones sobre estas bases.

= No necesita ningún tipo de conocimiento previo para realizar la clasificación.

}

Explicación: Transparencias 2 y 3 del tema 8.

Transparencia 2: El aprendizaje bayesiano usado como clasificador puede obtener probabilidades de pertenecer a cada clase.

Transparencia 3: El aprendizaje bayesiano busca la hipótesis "h" (de entre todas las "H" posibles) más probable si hemos observado una serie de de datos.

83. ¿Qué tipo de aprendizaje tiene una medida de lo bien o lo mal que está funcionando el algoritmo, pero no sabemos exactamente lo que falla?{

~ Aprendizaje supervisado.

~ Aprendizaje no supervisado.

= Aprendizaje por refuerzo.

}

Respuesta: Aprendizaje por refuerzo, diapositiva 5 de Introducción al aprendizaje automático.

84. En cuanto al aprendizaje bayesiano...  {

~ Está basado en el teorema de Bayes, y es una técnica similar a las redes bayesianas consiguiendo mismos resultados

~ Está basado en el teorema de Bayes, y se usa exclusivamente como clasificador que obtiene probabilidades de pertenecer a cada clase.

= Permite combinar los datos de ejemplo con reconocimiento a priori.

Explicación: Mirando la definición del aprendizaje bayesiano comprobamos que está basado en el teorema de Bayes, aunque tiene posibilidad de construir representaciones más complejas que las redes bayesianas, y no está limitado exclusivamente a usarse como clasificador, aunque de buenos resultados. Por tanto las dos primeras quedan descartadas. La tercera respuesta es pura definición.

Tema 8 - Pág. 2

85. En un ejemplo clasificador de textos, tenemos que clasificar un texto en una categoría predefinida dado cualquier conjunto de palabras de un texto (atributos), en las simplificaciones podemos afirmar que: {

= La posición de las palabras en el texto importa o no dependiendo del objetivo.

~ Nunca importa la posición de las palabras en el texto, solo si están presentes.  
  
~ No se aplica ninguna simplificación en los clasificadores de textos.

}

Explicación: Como en el ejemplo visto en clase se consideró irrelevante la ubicación de las palabras pero, por ejemplo, para determinar la ubicación de una web en sus resultados, Google valora mas las palabras que aparecen al principio del texto que las que se encuentran al final.

Algunas razones para ello son:

- Se considera que las primeras palabras son las que mejor describen el contenido de un texto.

- Las primeras palabras son vistas por muchos mas lectores.

- A partir de la lectura de las primeras palabras de una web, decidimos si leeremos resto del documento o seguiremos buscando.