

TEMA 8

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.

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)

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

{

= $h_{\text{MAP}} = \text{argmax} P(h|D)$

~ $h_{\text{MAP}} = \text{argmax} P(D|h)$

~ 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 $P(h) = \text{cte}$ y máxima verosimilitud.

¿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.

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.

Considerando 3 variables $A=1$, $B=5$, $C=2$, el resultado devuelto por $\text{argmax}(A, B, C)$ es :{

~ 5

= B

~ C

}

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

~ MAP implica un máximo a priori.

= ML implica máxima verosimilitud, debido a que $P(h) = \text{cte.}$

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

}

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.

}

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.

}

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 sobreentrenamiento.

~ Para realizar el entrenamiento en métodos de aprendizaje supervisado

}

En el aprendizaje supervisado, la regresión: {

~ Calcula una función continua.

= Estima una función continua.

~ Ninguna de las anteriores.

}

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

}

¿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

}

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).

}

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.

}

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

= Aprendizaje supervisado.

~ Aprendizaje no supervisado.

~ Aprendizaje por refuerzo.

}

¿En que consiste el aprendizaje automatico? :{

~ Como una computadora no puede mejorar con la experiencia, 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.

}

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.

}

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.

}

Dada esta definicion: "Es un método de aprendizaje Automático 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

}

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

}

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.

}

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

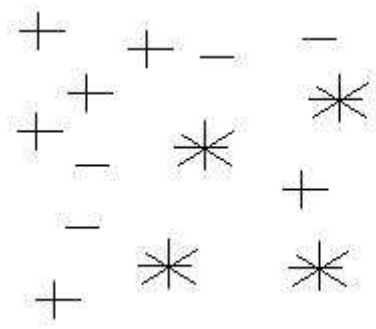
~ Mediante la validación cruzada.

= A través del error de validación.

~ No tiene solución.

}

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.

}

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).

}

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.

}

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

~ Entrenamiento

= Cluster

~ Validación

}

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

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

En el clasificador bayesiano:

{

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

~ Suponemos ejemplos caracterizados como tuplas de atributos $\langle a_1, a_2, \dots, a_n \rangle$

= Las dos son correctas.

}

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

}

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

}

Respecto al clasificador bayesiano 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

}

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.

}

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

= Aprendizaje no supervisado

~ Aprendizaje supervisado

~ Aprendizaje por refuerzo

}

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 $|\text{Voc}|$ en el denominador del cálculo de la probabilidad de una palabra en una categoría dada: $P(w_k|c_j) = (n_k+1)/(n+|\text{Voc}|)$: {

~ 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. }

Estamos implementando un modelo de aprendizaje para guiar a nuestro robot autónomo "Emilio" en un entorno laberíntico mediante sucesivas pruebas a base de prueba/error; y utilizando 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
}
```

El aprendizaje de bayesiano usado como clasificador: {

```
~ puede obtener probabilidades de no pertenecer a ninguna clase
= puede obtener probabilidades de pertenecer a cada clase.
~ ninguna de las anteriores
}
```

¿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".
}
```

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

DÍA	TEMP.	DIR. VIENTO	CIELO	PRESIÓN	Tiempo
1	≤ 0	Sur	Nuboso	Subiendo	Sol
2	> 0	Oeste	Claro	Estable	Sol
3	> 0	Norte	Claro	Subiendo	Sol
4	> 0	Norte	Claro	Bajando	Lluvia
5	> 0	Oeste	Nuboso	Bajando	Lluvia
6	≤ 0	Norte	Nuboso	Bajando	Nieve
7	> 0	Sur	Nuboso	Estable	Lluvia
8	> 0	Sur	Claro	Subiendo	Sol
9	≤ 0	Este	Nuboso	Bajando	Nieve
10	≤ 0	Sur	Claro	Estable	Sol

Podemos decir que la predicción del tiempo para

<presión=Subiendo,cielo=Nuboso>

sería: {

~ Lluvia

~ Nieve

= Sol

}

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 impartía la teoría era titular o asociado y la base de conocimiento de dicha asignatura (Programación, sistemas o teoría 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	Teoría de la información	Primero	Asociado	Suspenso
E	Programación	Primero	Titular	Sobresaliente
F	Sistemas	segundo	Titular	Aprobado
G	Teoría de la información	segundo	Asociado	Notable
H	Programación	segundo	Asociado	Sobresaliente
I	Programación	segundo	Titular	Suspenso
J	Sistemas	Primero	Titular	Notable
K	Teoría 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 teoría 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.

$$P(\text{Sobresaliente}) = 4/12 = 1/3$$

$$P(\text{BC} = \text{Programación} \mid \text{Sobresaliente}) = 1$$

$$P(\text{P} = \text{Asociado} \mid \text{Sobresaliente}) = 1/2$$

$$P(\text{C} = \text{Segundo} \mid \text{Sobresaliente}) = 1/4$$

$$P(\text{Programación, Asociado, Segundo} \mid \text{Sobresaliente}) = 1/3 * 1 * 1/2 * 1/4 = 1/24$$

$$P(\text{Notable}) = 3/12 = 1/4$$

$$P(\text{BC} = \text{Programación} \mid \text{Notable}) = 0$$

$$P(\text{Programación, Asociado, Segundo} \mid \text{Notable}) = 0$$

$$P(\text{Aprobado}) = 3/12 = 1/4$$

$$P(\text{BC} = \text{Programación} \mid \text{Aprobado}) = 0$$

$$P(\text{Programación, Asociado, Segundo} \mid \text{Aprobado}) = 0$$

$$P(\text{Suspenso}) = 2/12 = 1/6$$

$$P(\text{BC} = \text{Programación} \mid \text{Suspenso}) = 1/2$$

$$P(\text{P} = \text{Asociado} \mid \text{Suspenso}) = 1/2$$

$$P(\text{C} = \text{Segundo} \mid \text{Suspenso}) = 1/2$$

$$P(\text{Programación, Asociado, Segundo} \mid \text{Suspenso}) = 1/6 * 1/2 * 1/2 * 1/2 = 1/48$$

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

= Aprendizaje supervisado

~ Aprendizaje no supervisado

~ Aprendizaje por refuerzo

}

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

{

~ $P(c_j) = \text{nº de ejemplos no etiquetados con } c_j / \text{nº de ejemplos etiquetados con } c_j$

= $P(c_j) = \text{nº de ejemplos etiquetados con } c_j / \text{nº total de ejemplos}$

~ $P(c_j) = \text{nº total de ejemplos} / \text{nº de ejemplos etiquetados con } c_j$

}

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.

}

La **minería de datos** intenta descubrir patrones en grandes volúmenes de conjuntos de datos. Utiliza los métodos de la inteligencia artificial, aprendizaje automático, estadística 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.

}

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.

}

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 predicciones correctas de la llegada del autobus para <pasajeros = muchos, tiempo = sol> utilizando $P(c_i) \prod_j P(a_j / c_i)$?

{

~ $c_i = \text{puntual} \rightarrow 1/9, c_i = \text{temprano} \rightarrow 0, c_i = \text{retrasado} \rightarrow 1/18$

= $c_i = \text{puntual} \rightarrow 1/12, c_i = \text{temprano} \rightarrow 0, c_i = \text{retrasado} \rightarrow 1/9$

~ $c_i = \text{puntual} \rightarrow 1/12, c_i = \text{temprano} \rightarrow 1, c_i = \text{retrasado} \rightarrow 1/9$

}

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.

}

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.

}

Dadas las variables $X=9$, $Y=3$, $Z=5$, la solución a $\max(X,Y,Z)$ y a $\operatorname{argmax}(X,Y,Z)$ es :{
= 9 y X respectivamente.

~ X y 9 respectivamente.

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

}

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.

}

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.

}

Después de realizar el cálculo del siguiente caso de predicción meteorológica, obtenemos $P(T<0, \text{Viento}=\text{Norte}|\text{Lluvia})=0$ ¿Qué podemos determinar? {

~ No puede ser 0, no es una probabilidad.

~ Es seguro que llueva.

= No lloverá.

}

Cuál de las siguientes afirmaciones acerca del aprendizaje bayesiano es falsa:

{

~ Permite combinar los datos de ejemplo con conocimiento a priori

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

= Esta basado en el teorema $\rightarrow P(h|D) = P(D) / P(D|h)P(h)$

}

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: {

~ En este caso si importa la posición de las palabras en el texto.

= No importa la posición de las palabras en el texto, solo si están presentes.

~ No se aplica ninguna simplificación en los clasificadores de textos.

}

Google completa la cadena que introduces en la barra del buscador, utiliza el modelo oculto de markov que es un modelo estadístico 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 ningún tipo de aprendizaje automático.

= Se basa en el tipo de aprendizaje supervisado.

~ Se basa en el tipo de aprendizaje no supervisado.

}

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.

}

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.

}

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.

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> {

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

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

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

}

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

~ $\text{n}^\circ \text{ total de ejemplos} / \text{n}^\circ \text{ de ejemplos de la clase actual}$

~ $\text{n}^\circ \text{ de ejemplos con atributos } a_1, a_2, \dots, a_n / \text{n}^\circ \text{ de ejemplos de la clase actual}$

= $\text{n}^\circ \text{ de ejemplos de la clase actual} / \text{n}^\circ \text{ total de ejemplos}$

}

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

}

¿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 del conocimiento especializado para resolver problemas como un especialista humano.

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

}

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.

}

¿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

}

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

~ No es necesario conocimiento a priori.

= Es necesario conocimiento a priori.

~ Tiene un bajo costo computacional.

}

¿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.

}

¿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

}

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 hipótesis 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.

}

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 hipótesis 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.

}

El aprendizaje bayesiano se basa en el teorema de Bayes, el cual es: {

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

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

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

}

Respecto al aprendizaje bayesiano indica la opción correcta: {

~ En la fase 3 hallamos la estimación MAP.

~ En la fase 2 calculamos $docs_j$, documentos de la clase c_j .

= Ninguna de las anteriores.

}

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.

}

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

Vuelta	Temperatura	Estado	Presión	Tiempo
1	30	Seco	Media	Soleado
2	30	Seco	Alta	Soleado
3	29	Seco	Alta	Nublado
4	30	Seco	Alta	Soleado
5	30	Seco	Media	Lluvia
6	29	Mojado	Baja	Lluvia
7	29	Mojado	Baja	Lluvia
8	30	Mojado	Baja	Nublado
9	31	Mojado	Baja	Soleado
10	30	Mojado	Media	Soleado
11	31	Mojado	Baja	Soleado
12	32	Seco	Alta	Soleado

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

~ 0.58

= 0

~ 0.25

}

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.

}

En la fase de clasificación: {

= Dado un documento $x=w_1, w_2, \dots, w_n$, nos quedamos con las posiciones de palabras que están contenidas en Voc y devolvemos la estimación map.

~ Dado un documento $x=w_1, w_2, \dots, w_n$, nos quedamos con las posiciones de palabras que no están contenidas en Voc y devolvemos la estimación map

~ Ambas son incorrectas.

}

Cuáles de los siguientes afirmaciones acerca de problemas de la aprendizaje bayesiano NO son correctas?

{

~ Necesidad de un conocimiento a priori.

~ Costo computacional alto.

= No hay la posibilidad de construir representaciones más complejas.

}

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

Cuáles son las predicciones correctas de la llegada del autobus para <tráfico = normal, tiempo = lluvia> utilizando el clasificador bayesiano ($P(c_i) \prod_j P(a_j/c_i)$)?

{

- ~ $c_i = \text{puntual} \rightarrow 5/24, c_i = \text{temprano} \rightarrow 0, c_i = \text{retrasado} \rightarrow 1/9$
- ~ $c_i = \text{puntual} \rightarrow 3/24, c_i = \text{temprano} \rightarrow 0, c_i = \text{retrasado} \rightarrow 1/8$
- = $c_i = \text{puntual} \rightarrow 3/32, c_i = \text{temprano} \rightarrow 0, c_i = \text{retrasado} \rightarrow 1/8$

}

En una red Bayesiana existe Maxima Verosimilitud (ML) cuando {

- ~ $P(h) = \text{no cte}$
- = $P(h) = \text{cte}$
- ~ No depende de $P(h)$

}

¿Qué presupone el clasificador bayesiano „naive“? {

- ~ Que todos los atributos tienen que ser dependientes uno del otro
- ~ Que todos los atributo siempre tienen dos nodos padres
- = Que todos los atributos tienen que ser independientes

}

La siguiente fórmula es la base del aprendizaje bayesiano:

$$h_MAP = \text{argmax} P(h | D)$$

¿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.

}

Podemos calcular la probabilidad de los atributos condicionada a las diferentes

categorías $P(a_1, a_2, \dots, a_n | c_i)$ con la aproximación $\prod_{j=1}^n P(a_j | c_i)$ considerando: {

= Los atributos son independientes entre si.

~ La probabilidad de las clases se suponen las mismas.

~ Ninguna de las anteriores.

}

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

{

~Porque las clases se relacionan con atributos independientes.

=Para que cuando la probabilidad dé 0 normalmente, esté 0 no sea muy significativo.

~Ninguna de las dos es cierta.

}

¿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.

}