uc3m | Universidad Carlos III de Madrid

Ejercicios Modelos de Markov

Modelos Ocultos de Markov (HMM)

Inteligencia Artificial Colmenarejo

Curso 2022-2023

Resumen teórico

- Vemos los Modelos de Markov como un caso particular de Redes Bayesianas que representan una secuencia de valores para una variable (o variables), donde dichos valores no son independientes entre sí.
- Las variables que representan el mismo objeto en diferentes momentos de la secuencia se indican mediante un subíndice (X_i).
- Sólo consideramos Modelos de Markov estacionarios de primer orden $(i|i-1)$. Por lo tanto, se describen con dos distribuciones de probabilidad:
  - P(X_0) Probabilidad a priori para la variable inicial X.
  - P($X_i\,|\,X_{i-1}$) Probabilidad de transición para X, estacionaria en la secuencia.
- En los Modelos Ocultos de Markov (HMM), tenemos también variables Evidencia E que dependen en un estado oculto (variables latentes X), y son condicionalmente independientes del resto de la estructura (dado X). Por lo tanto, el HMM introduce otra distribución de probabilidad condicional:
  - $P(E_i|X_i)$ Probabilidad de emisión para E, estacionaria durante la secuencia.
- La inferencia se realiza el método exacto genérico para las RB, aunque por su estructura se pueden aplicar métodos de sustitución de variables en algunos casos (inferencia hacia adelante).

Para resolver un HMM hay que realizar los siguientes pasos:

1. Modelar el HMM: un HMM se define mediante el estado oculto del problema (O_T) y su observación (A_T).
2. Modelar la representación de la red bayesiana: dibujar la red bayesiana del problema.
3. Definir las distribuciones de probabilidad que afectan al problema y colocarlas en la red bayesiana.
4. Definir la tarea de inferencia a resolver.

Enunciado

Con un radar se obtiene cada pocos segundos una observación, compuesta por un conjunto de atributos sobre la posición y la velocidad de un caza, los cuales de forma genérica se denotarán por O_T, siendo T un instante arbitrario.

Se desea construir un sistema utilizando un Modelo Oculto de Markov (HMM) para determinar si el caza supone una amenaza inminente o no.

Preguntas:

- ¿Cómo se realizaría el razonamiento necesario para obtener la solución?

Solución

¿Cómo se realizaría el razonamiento necesario para obtener la solución?

Enunciado: Con un radar se obtiene cada pocos segundos una observación, compuesta por un conjunto de atributos sobre la posición y la velocidad de un caza, los cuales de forma genérica se denotarán por O_T, siendo T un instante arbitrario.

Solución

¿Cómo se realizaría el razonamiento necesario para obtener la solución?

1. **Modelar el HMM definiendo el estado oculto y la observación**
   - Observación O_T:
   - Estado oculto A_T:
2. **Modelar la representación de la red bayesiana**

Enunciado: Con un radar se obtiene cada pocos segundos una observación, compuesta por un conjunto de atributos sobre la posición y la velocidad de un caza, los cuales de forma genérica se denotarán por O_T, siendo T un instante arbitrario.

Solución

¿Cómo se realizaría el razonamiento necesario para obtener la solución?

1. **Modelar el HMM definiendo el estado oculto y la observación**
   - Observación O_T: El conjunto de valores que proporciona el radar, que incluyen posición y velocidad.
   - Estado oculto A_T: Representa si el caza es una amenaza o no (a,¬a) en el instante t.

2. **Modelar la representación de la red bayesiana**

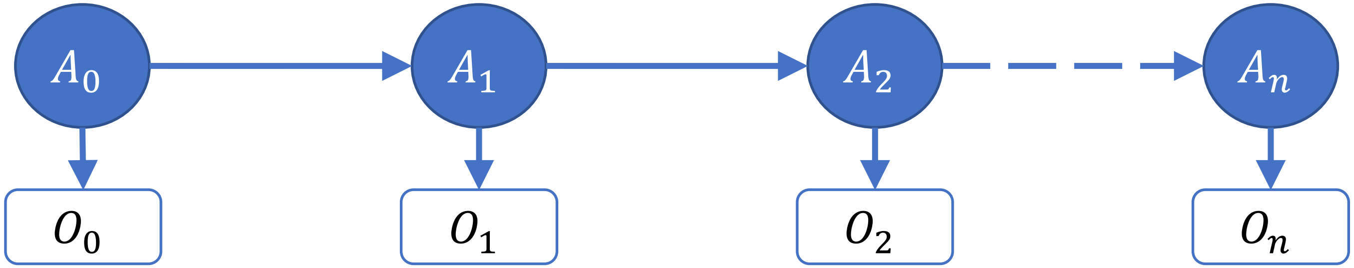
Enunciado: Con un radar se obtiene cada pocos segundos una observación, compuesta por un conjunto de atributos sobre la posición y la velocidad de un caza, los cuales de forma genérica se denotarán por O_T, siendo T un instante arbitrario.

Solución

¿Cómo se realizaría el razonamiento necesario para obtener la solución?

1. **Modelar el HMM definiendo el estado oculto y la observación**
   - Observación O_T: El conjunto de valores que proporciona el radar, que incluyen posición y velocidad.
   - Estado oculto A_T: Representa si el caza es una amenaza o no (a,¬a) en el instante t.

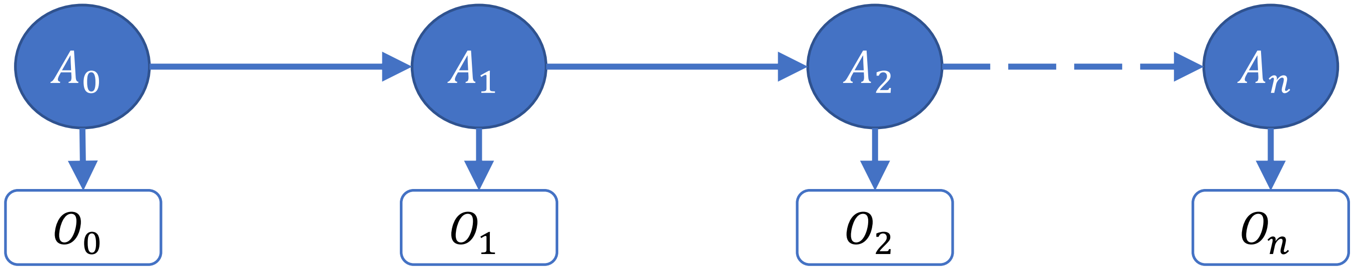
2. **Modelar la representación de la red bayesiana**



Solución

3. Definir las distribuciones de probabilidad

Enunciado: Con un radar se obtiene cada pocos segundos una observación, compuesta por un conjunto de atributos sobre la posición y la velocidad de un caza, los cuales de forma genérica se denotarán por O_T, siendo T un instante arbitrario.

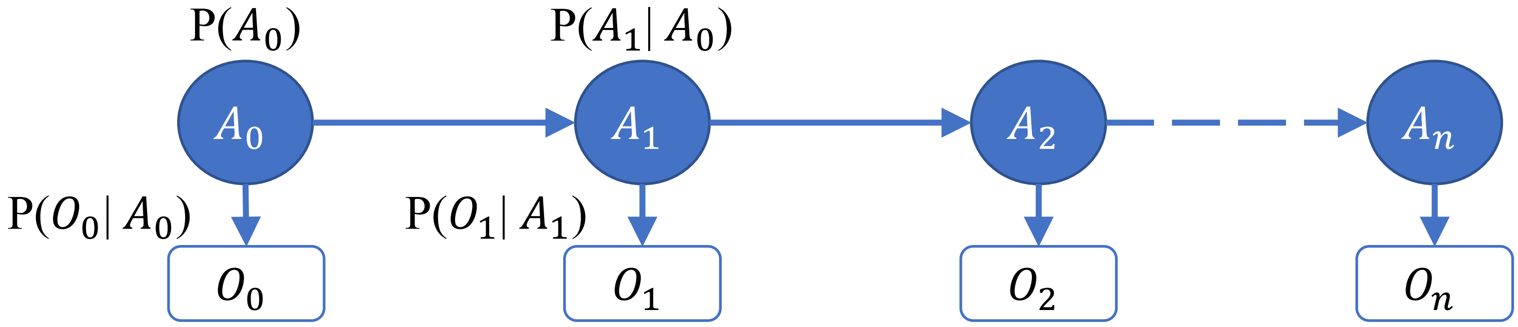


4. Definir la tarea de inferencia a resolver

Solución

3. Definir las distribuciones de probabilidad

- La distribución P(A_0), que modelaría la probabilidad inicial de que el caza suponga una amenaza.
- La distribución P($A_{T+1} \mid A_T$) nos permitiría modelar la evolución temporal del estado del caza, es decir, la probabilidad de que en un instante suponga una amenaza o no dependiendo del estado anterior.
- La distribución P($O_T \mid A_T$) permite modelar las observaciones que obtiene el radar de acuerdo a si el caza constituye una amenaza o no.

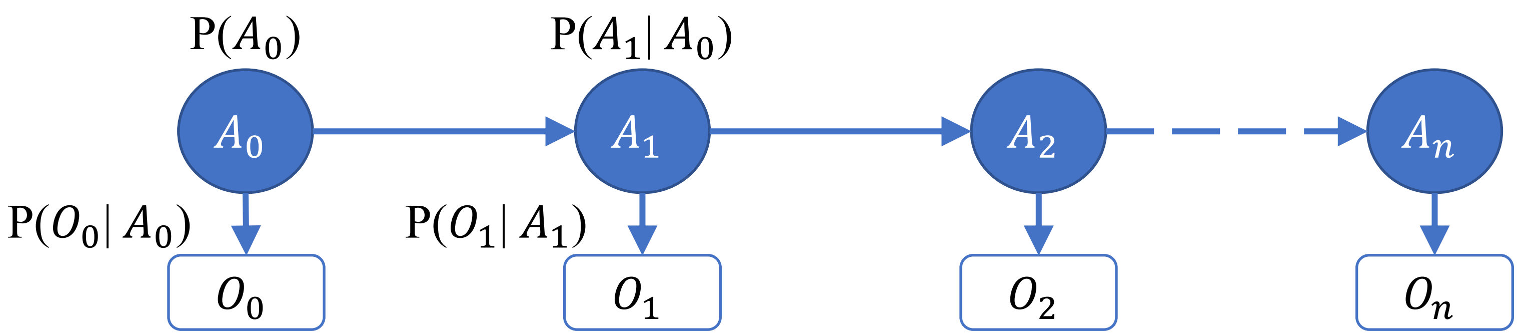


4. Definir la tarea de inferencia a resolver

Solución

3. Definir las distribuciones de probabilidad

- La distribución P(A_0), que modelaría la probabilidad inicial de que el caza suponga una amenaza.
- La distribución P($A_{T+1} | A_T$) nos permitiría modelar la evolución temporal del estado del caza, es decir, la probabilidad de que en un instante suponga una amenaza o no dependiendo del estado anterior.
- La distribución P($O_T | A_T$) permite modelar las observaciones que obtiene el radar de acuerdo a si el caza constituye una amenaza o no.



4. Definir la tarea de inferencia a resolver

Argmax$_{a,\neg a}$ P(A_T | $O_1 \ldots O_T$) = max[P(a | $O_1 \ldots O_T$), P(¬a| $O_1 \ldots O_T$)]

Es decir, calcularíamos las probabilidades para el tiempo T de que la amenaza sea “a”, o “¬a”. Del máximo de ambas e obtiene dicho argumento “a” o “¬a”.

Ejercicio 2

Enunciado

¿Cómo podría aplicar HMMs (Modelos Ocultos de Markov) para decidir si un alumno está entendiendo o no una asignatura?

Describa todos los componentes del HMM correspondiente, asignando razonadamente probabilidades numéricas.

Solución

Como siempre, hay que seguir los pasos:

1. Modelar el HMM: un HMM se define mediante el estado oculto del problema (O_T) y su observación (A_T)
2. Modelar la representación de la red bayesiana: dibujar la red bayesiana del problema
3. Definir las distribuciones de probabilidad que afectan al problema y colocarlas en la red bayesiana
4. Definir la tarea de inferencia a resolver

1. Modelar el HMM: un HMM se define mediante el estado oculto del problema (O_T) y su observación (A_T)

Solución

Como siempre, hay que seguir los pasos:

1. Modelar el HMM: un HMM se define mediante el estado oculto del problema (O_T) y su observación (A_T)
2. Modelar la representación de la red bayesiana: dibujar la red bayesiana del problema
3. Definir las distribuciones de probabilidad que afectan al problema y colocarlas en la red bayesiana
4. Definir la tarea de inferencia a resolver

1. Modelar el HMM: un HMM se define mediante el estado oculto del problema (O_T) y su observación (A_T)

En este caso no nos dejan excesivamente clara la observación, por lo que hay que definir una.

Sabemos que el estado no observable es el grado de comprensión de la asignatura, ya lo definamos como una variable booleana o con más valores discretos.

Ejercicio 2

Solución

Además, al usar HMM, debe existir una relación temporal entre el nivel de comprensión en los momentos sucesivos. Es decir, modelar el hecho de que un alumno nunca pasa en un solo paso de entender completamente la asignatura a no entender nada, o viceversa. Es decir, ambas variables no son independientes. Tenemos que hacer la suposición (fundada) de que sí que basta con especificar la dependencia entre un paso y el anterior (modelo de Markov de orden 1).

La evidencia podrían ser los sucesivos exámenes de la asignatura, dado que la nota es conocida una vez realizados.

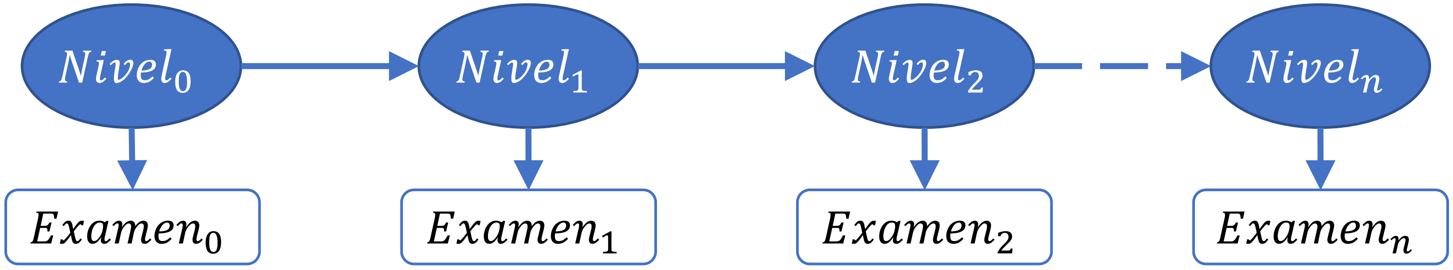
Solución

Además, al usar HMM, debe existir una relación temporal entre el nivel de comprensión en los momentos sucesivos. Es decir, modelar el hecho de que un alumno nunca pasa en un solo paso de entender completamente la asignatura a no entender nada, o viceversa. Es decir, ambas variables no son independientes. Tenemos que hacer la suposición (fundada) de que sí que basta con especificar la dependencia entre un paso y el anterior (modelo de Markov de orden 1).

La evidencia podrían ser los sucesivos exámenes de la asignatura, dado que la nota es conocida una vez realizados.

2. Modelar la representación de la red bayesiana: dibujar la red bayesiana del problema

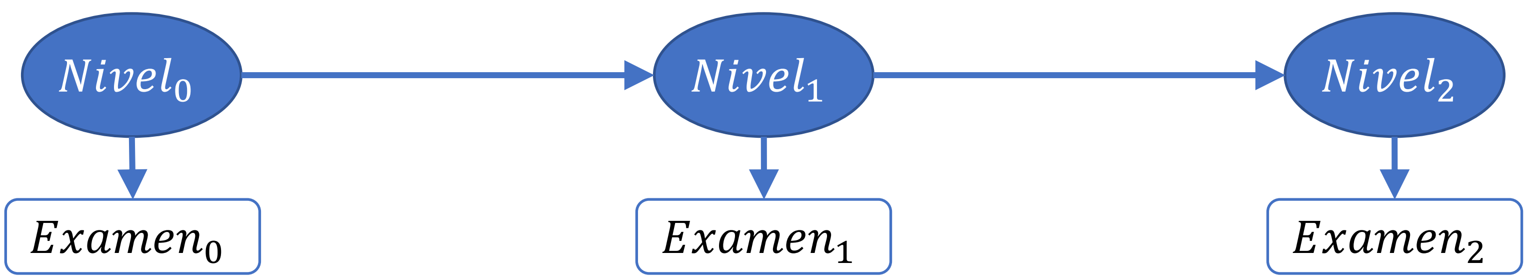
Podemos representar esto con la figura siguiente. La secuencia de exámenes es la que define la secuencia temporal (no son tiempos fijos).



Ejercicio 2

Solución

3. Definir las distribuciones de probabilidad que afectan al problema y colocarlas en la red bayesiana

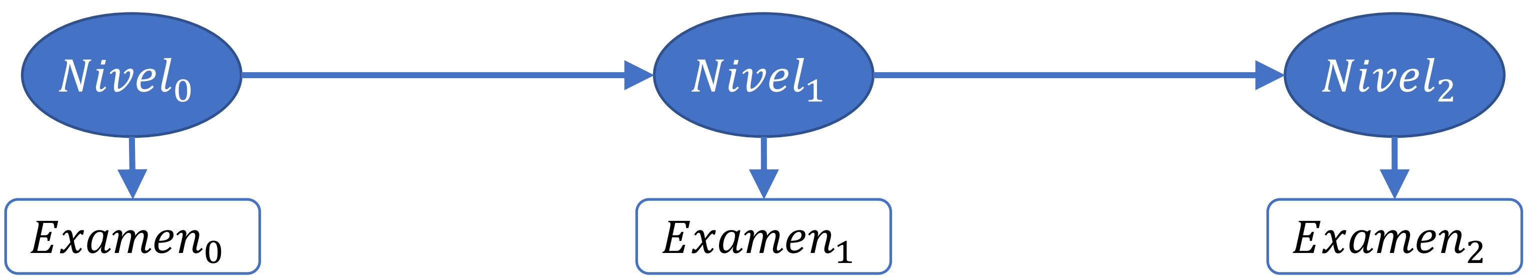


Ejercicio 2

Solución

3. Definir las distribuciones de probabilidad que afectan al problema y colocarlas en la red bayesiana

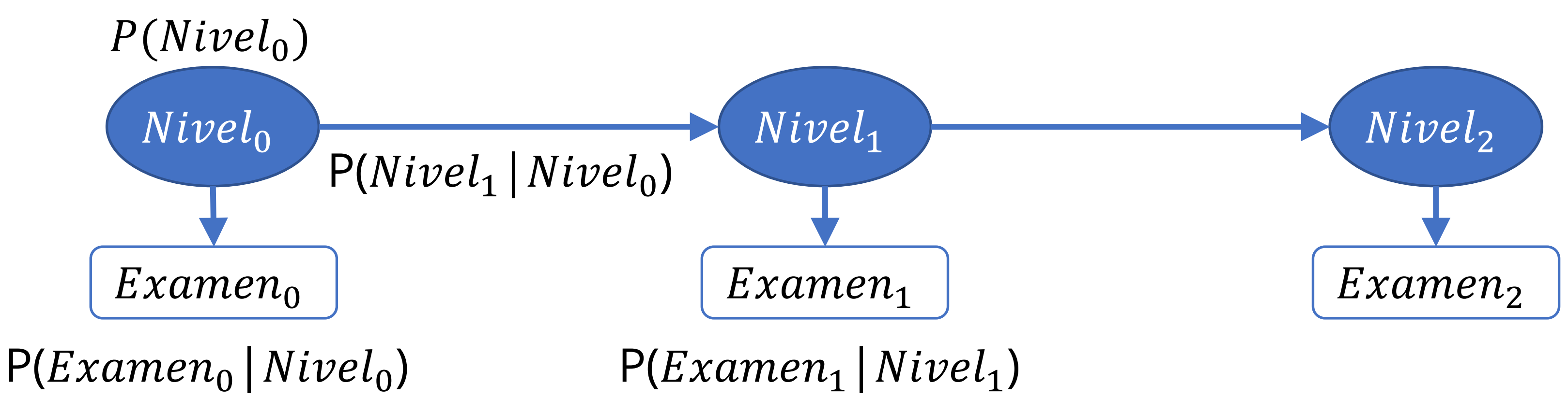
- $P(Nivel_0)$: La probabilidad a priori del estado oculto.
- P($Nivel_i \,|\, Nivel_{i-1}$) para i > 0. En un caso real habría que fundar estos datos en la experiencia. Para este ejemplo vamos a asignarlos simplemente teniendo en cuenta que hay mayor probabilidad de mejorar que de empeorar.
- P($Examen_i \,|\, Nivel_i$). Aquí vamos a entender que si se tiene nivel N se aprueba casi siempre, mientras que con los otros niveles hay mayor incertidumbre.



Solución

3. Definir las distribuciones de probabilidad que afectan al problema y colocarlas en la red bayesiana

- $P(Nivel_0)$: La probabilidad a priori del estado oculto.
- P($Nivel_i \,|\, Nivel_{i-1}$) para i > 0. En un caso real habría que fundar estos datos en la experiencia. Para este ejemplo vamos a asignarlos simplemente teniendo en cuenta que hay mayor probabilidad de mejorar que de empeorar.
- P($Examen_i \,|\, Nivel_i$). Aquí vamos a entender que si se tiene nivel N se aprueba casi siempre, mientras que con los otros niveles hay mayor incertidumbre.



Solución

Como nos piden dar valores, tenemos que fijar el dominios de cada variables.

Por simplificar las tablas podríamos definir:

- Nivel ∊ {I, S, N} (I es nivel insuficiente, S suficiente y N notable)
- Examen ∊ {T , F } (T es que el examen se aprueba).

Solución

Como nos piden dar valores, tenemos que fijar el dominios de cada variables.

Por simplificar las tablas podríamos definir:

- Nivel ∈ {I, S, N} (I es nivel insuficiente, S suficiente y N notable)
- Examen ∈ {T , F } (T es que el examen se aprueba).

Suponiendo un modelo estacionario, hay que definir:

- $P(Nivel_0)$: {0.75(I), 0.25(S), 0(N)}, es más probable tener un nivel bajo al principio.
- y las dos tablas de probabilidad condicional

	Nivel$_i$		
Nivel$_{i-1}$	I	S	N
I	0.50	0.30	0.20
S	0.20	0.50	0.30
N	0.0	0.20	0.80

	Examen$_i$	
Nivel$_i$	T	F
I	0.30	0.70
S	0.70	0.30
N	0.90	0.10

Ejercicio 2

Solución

La tarea de predecir el nivel del alumno en el paso i sería el cálculo:
$P(Nivel_i \,|\, Examen_{i-1}, Examen_{i-2}, \cdots, Examen_0)$

Se podría calcular a partir de las notas de los exámenes anteriores, utilizando el procedimiento general para redes bayesianas, o el método de propagación hacia delante. En cualquier caso, sólo se utilizan las variables del pasado (las del futuro se cancelarán).

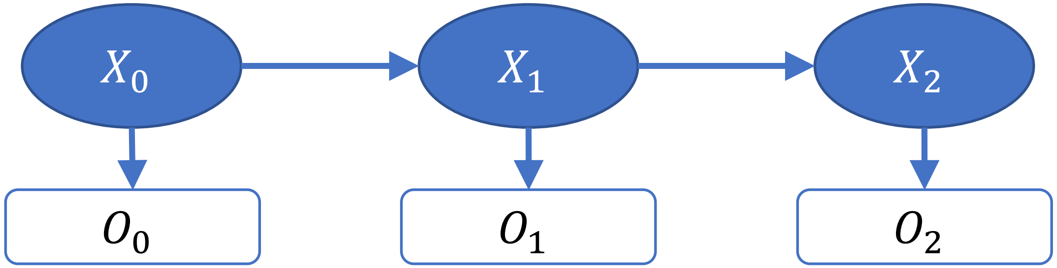
Enunciado

A veces uno tiene un constipado (C) que le hace estornudar (E). También puede tener alergia (A) y esto también le puede hacer estornudar. Otras veces uno se siente bien (B) y no estornuda (¬e).

Se decide modelar esto como un HMM donde:

- Los estados X_t representan el estado de la persona:
    - {bien (B), alergia (A), constipado (C)}
- Las observaciones O_t representan los síntomas observables:
    - {e, ¬e}

La representación del HMM es la siguiente:



Enunciado

Dadas las probabilidades iniciales y distribuciones de probabilidad del problema:

P(X_1) = (1, 0, 0) (i.e. P(X_1 = B) = 1)

P(X_{t+1}/X_t)

$\boldsymbol{X_t}$	$\mathbf{P}(\boldsymbol{X_{t+1} = B/X_t})$	$\mathbf{P}(\boldsymbol{X_{t+1} = A/X_t})$	$\mathbf{P}(\boldsymbol{X_{t+1} = C/X_t})$
B	0.7	0.2	0.1
A	0.6	0.3	0.1
C	0.2	0.2	0.6

P(O_t / X_t)

$\boldsymbol{X_t}$	$\mathbf{P}(\boldsymbol{O_t = e/X_t})$	$\mathbf{P}(\boldsymbol{O_t = \neg e/X_t})$
B	0.1	0.9
A	0.8	0.2
C	0.7	0.3

Preguntas:

1. ¿Cuál es la probabilidad de que uno esté bien (B) mañana (día 2)?
2. ¿Cuál es la probabilidad de la secuencia B,C,C,B en los primeros 4 días?
3. ¿Cuál es la probabilidad de que en el día 1 se observe no estornudo (¬s) y en el día 2 se observe estornudo (e)?

Solución

¿Cuál es la probabilidad de que uno esté bien (B) mañana (día 2)?

¿Cuál es la probabilidad de la secuencia B,C,C,B en los primeros 4 días?

¿Cuál es la probabilidad de que en el día 1 se observe no estornudo (¬e) y en el día 2 se observe estornudo (e)?

Solución

¿Cuál es la probabilidad de que uno esté bien (B) mañana (día 2)?

$$\Pr(X_2 = B) = \Pr(X_2 = B|X_1 = B)\Pr(X_1 = B) = 0.7(1) = 0.7$$

¿Cuál es la probabilidad de la secuencia B,C,C,B en los primeros 4 días?

¿Cuál es la probabilidad de que en el día 1 se observe no estornudo (¬e) y en el día 2 se observe estornudo (e)?

Solución

¿Cuál es la probabilidad de que uno esté bien (B) mañana (día 2)?

$$\Pr(X_2 = B) = \Pr(X_2 = B | X_1 = B) \Pr(X_1 = B) = 0.7(1) = 0.7$$

¿Cuál es la probabilidad de la secuencia B,C,C,B en los primeros 4 días?

$$\begin{aligned}\Pr(B_1, C_2, C_3, B_4) &= \Pr(B_4 | C_3) \Pr(C_3 | C_2) \Pr(C_2 | B_1) \Pr(B_1) \\ &= 0.2\,(0.6)\,(0.1)\,(1) = 0.012\end{aligned}$$

¿Cuál es la probabilidad de que en el día 1 se observe no estornudo (¬e) y en el día 2 se observe estornudo (e)?

Solución

¿Cuál es la probabilidad de que uno esté bien (B) mañana (día 2)?

$$\Pr(X_2 = B) = \Pr(X_2 = B | X_1 = B) \Pr(X_1 = B) = 0.7(1) = 0.7$$

¿Cuál es la probabilidad de la secuencia B,C,C,B en los primeros 4 días?

$$\Pr(B_1, C_2, C_3, B_4) = \Pr(B_4 | C_3) \Pr(C_3 | C_2) \Pr(C_2 | B_1) \Pr(B_1)$$
$$= 0.2\,(0.6)\,(0.1)\,(1) = 0.012$$

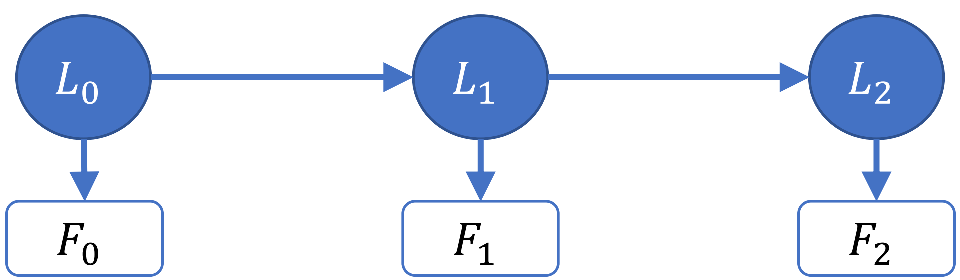
¿Cuál es la probabilidad de que en el día 1 se observe no estornudo (¬e) y en el día 2 se observe estornudo (e)?

$$\Pr(\neg e_2, e_2) = \sum_{x_2} \Pr(e|X) \Pr(X|B) \Pr(\neg e|B) \Pr(B)$$
$$= \Pr(e_2|W_2) \Pr(B_2|B_1) \Pr(\neg e_1|B_1) \Pr(B_1) +$$
$$\Pr(e_2|A_2)\ \Pr(A_2|B_1)\ \Pr(\neg e_1|B_1) \Pr(B_1) +$$
$$\Pr(e_2|C_2)\ \Pr(C_2|B_1)\ \Pr(\neg e_1|B_1) \Pr(B_1)$$
$$= 0.1(0.7)(0.9)(1) + 0.8(0.2)(0.9)(1) + 0.7(0.1)(0.9)(1) = 0.27$$

Enunciado

Queremos construir un sistema para reconocimiento automático del lenguaje. Concretamente, queremos comprender las palabras que corresponden a ciertas señales acústicas: cada letra corresponde en teoría a una de estas señales. Para simplificar consideraremos un lenguaje en el que hay sólo tres letras (a, r, t), y también las palabras sólo pueden tener tres letras de longitud. El proceso se puede modelar con un HMM con una sola variable oculta que representa la letra, y las observaciones la señal acústica, en forma de un fonema extraído de la señal.

- Los estados L_t representan la letra {a, r , t}
- Las observaciones F_t representan el fonema {a, r , t}



Enunciado

Se conocen las siguientes probabilidades a priori y probabilidades de transición condicional, y que el modelo es estacionario.

$$P(L_0) = (0.3, 0.3, 0.4)$$

$$P(L_{t+1}/L_t)$$

$\boldsymbol{L_t}$	$\boldsymbol{P(L_{t+1} = a/L_t)}$	$\boldsymbol{P(L_{t+1} = r/L_t)}$	$\boldsymbol{P(L_{t+1} = t/L_t)}$
a	0.0	0.7	0.3
r	0.5	0.0	0.5
t	1.0	0.0	0.0

$$P(F_t/L_t)$$

$\boldsymbol{L_t}$	$\boldsymbol{P(F_t = a/L_t)}$	$\boldsymbol{P(F_t = r/L_t)}$	$\boldsymbol{P(F_t = t/L_t)}$
a	0.9	0.1	0.0
r	0.1	0.8	0.1
t	0.0	0.2	0.8

Enunciado

Se pide:

1. ¿Cuál es la probabilidad de que la letra inicial sea a cuando el sistema escucha el fonema a? $P(L_0 = a/F_0 = a)$
2. ¿Cuál es la probabilidad de escuchar el fonema a en primera posición? $P(F_0 = a)$
3. ¿Cómo computaría la probabilidad de que la palabra emitida sea "art"?

Solución

$P(L_0) = (0.3, 0.3, 0.4)$

$P(L_{t+1}/L_t)$

L_t	$P(L_{t+1} = a/L_t)$	$P(L_{t+1} = r/L_t)$	$P(L_{t+1} = t/L_t)$
a	0.0	0.7	0.3
r	0.5	0.0	0.5
t	1.0	0.0	0.0

$P(F_t/L_t)$

L_t	$P(F_t = a/L_t)$	$P(F_t = r/L_t)$	$P(F_t = t/L_t)$
a	0.9	0.1	0.0
r	0.1	0.8	0.1
t	0.0	0.2	0.8

1. **¿Cuál es la probabilidad de que la letra inicial sea a cuando el sistema escucha el fonema a? $\boldsymbol{P(L_0 = a/F_0 = a)}$**

Para calcular $P(L_0 = a/F_0 = a)$ el método general requeriría hacer:

$$P(L_0 = a/F_0 = a) = \alpha\, P(L_0 = a, F_0 = a) =$$

$$= \alpha \sum_{l_{1i}} \sum_{f_{1i}} \dots P(L_0 = a, F_0 = a, L_1 = l_{1i}, F_1 = f_{1i}, \dots)$$

Como las variables para *i* > 0 no son antecesoras ni de consulta ni de evidencia, podemos ignorarlas (esos sumandos se cancelan, puesto que suman 1.0). Al ser solamente dos variables, nos queda una aplicación sencilla del teorema de Bayes:

$$P(L_0 = a/F_0 = a) = \alpha\, P(L_0 = a, F_0 = a) = \alpha\, P(L_0 = a) \cdot P(F_0 = a \mid L_0 = a)$$

$$P(L_0 = a/F_0 = a) = 1/P(F_0 = a) \cdot 0.3 \cdot 0.9$$

Solución

$P(L_0) = (0.3, 0.3, 0.4)$

$P(L_{t+1}/L_t)$

L_t	$P(L_{t+1} = a/L_t)$	$P(L_{t+1} = r/L_t)$	$P(L_{t+1} = t/L_t)$
a	0.0	0.7	0.3
r	0.5	0.0	0.5
t	1.0	0.0	0.0

$P(F_t/L_t)$

L_t	$P(F_t = a/L_t)$	$P(F_t = r/L_t)$	$P(F_t = t/L_t)$
a	0.9	0.1	0.0
r	0.1	0.8	0.1
t	0.0	0.2	0.8

2. ¿Cuál es la probabilidad de escuchar el fonema a en primera posición? $P(F_0 = a)$

Para calcular $P(F_0 = a)$ hay que usar la conjunta:

$$P(F_0 = a) = P(F_0 = a, L_0 = a) + P(F_0 = a, L_0 = r) + P(F_0 = a, L_0 = t)$$

Cada término se puede factorizar:

$$P(F_0 = a, L_0 = a) = P(L_0 = a) \cdot P(F_0 = a|L_0 = a) = 0.3 \cdot 0.9$$

$$P(F_0 = a, L_0 = r) = P(L_0 = r) \cdot P(F_0 = a\,|L_0 = r) = 0.3 \cdot 0.1$$

$$P(F_0 = a, L_0 = t) = P(L_0 = t) \cdot P(F_0 = a\;|L_0 = t) = 0.4 \cdot 0.0$$

Por lo tanto: $P(F_0 = a) = 0.3 \cdot 0.9 + 0.3 \cdot 0.1 = 0.3$

Este valor se puede sustituir en la pregunta anterior para determinar que:

$$P(L_0 = a|F_0 = a) = 0.9$$

Solución

$P(L_0) = (0.3, 0.3, 0.4)$

$P(L_{t+1}/L_t)$

L_t	$P(L_{t+1} = a/L_t)$	$P(L_{t+1} = r/L_t)$	$P(L_{t+1} = t/L_t)$
a	0.0	0.7	0.3
r	0.5	0.0	0.5
t	1.0	0.0	0.0

$P(F_t/L_t)$

L_t	$P(F_t = a/L_t)$	$P(F_t = r/L_t)$	$P(F_t = t/L_t)$
a	0.9	0.1	0.0
r	0.1	0.8	0.1
t	0.0	0.2	0.8

3. ¿Cómo computaría la probabilidad de que la palabra emitida sea "art"?

Para calcular la probabilidad de una determinada secuencia "art" en las variables ocultas, también se pueden ignorar las variables evidencia. Por ejemplo, para dos letras:

$P(L_0 = a, L_1 = r) = P(L_0 = a) \cdot P(L_1 = r|L_0 = a) \cdot$

$\sum_{f_0 \epsilon \{a,r,t\}} \sum_{f_1 \epsilon \{a,r,t\}} P(F_0 = f_0|L_0 = a) \cdot P(F_1 = f_1|L_1 = r)$

$\sum_{f_0 \epsilon \{a,r,t\}} P(F_0 = f_0|L_0 = a) \cdot \sum_{f_1 \epsilon \{a,r,t\}} P(F_1 = f_1|L_1 = r)$

¡Los sumatorios suman 1!

Luego, para dos letras: $P(L_0 = a, L_1 = r) = P(L_0 = a) \cdot P(L_1 = r|L_0 = a)$

Extendiendo para 3 variables: $P(L_0 = a, L_1 = r, L_2 = t) = P(L_0 = a) \cdot$

$P(L_1 = r|L_0 = a) \cdot P(L_2 = t|L_1 = r) = 0.3 \cdot 0.7 \cdot 0.5 = 0.105$

Ejercicios sin solución

A partir de este punto se presentan ejercicios sin solución que pueden ser resueltos por el alumno para apoyar el estudio de la asignatura.
Cualquier duda de los mismos será resuelta en la clase de en la clase de preparación del examen parcial 2 o tras dicha clase por mail al profesor de los ejercicios.

Enunciado

Se desea construir un sistema que asocie a cada palabra de una frase su categoría gramatical. Por ejemplo, para la frase: “El alumno aprobará Inteligencia Artificial” la salida deseada sería: “determinante nombre verbo nombre adjetivo”.

Preguntas:

- ¿Cuál de los modelos vistos en clase será el adecuado para construir este sistema?
- ¿Cómo se realizaría el razonamiento necesario para obtener la solución?

Ejercicio Tenis

Enunciado

Se dispone de la estadística de la figura sobre la conveniencia de jugar o no al tenis, en función de datos meteorológicos.
Construya un clasificador Naive Bayes que nos aconseje sobre hacer o no la actividad, usando datos de los primeros 10 días para predecir los últimos 4.

Preguntas

1. Construya Modelos de Markov que modelen de forma separada las cuatro variables meteorológicas. ¿Qué pasaría si quisiéramos tener en cuenta todas ellas en un solo modelo?
2. Construya un modelo oculto de Markov que permita deducir si hemos jugado al tenis un día del pasado, teniendo en cuenta como observación una sola de las variables meteorológicas (escoja la que prefiera).

Day	Weather	Temperature	Pressure	Wind	Cloud
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No