

Tareas de la asignatura “Métodos probabilistas”

Tema 2. Inferencia en redes bayesianas

▷ Fecha de entrega: 25 de enero de 2025.

Por favor, antes de entregar los ejercicios, lea de nuevo detenidamente las instrucciones que se encuentran en Ágora.

Ejercicio 2.1.

Sea la red bayesiana de 6 variables binarias dada por los enlaces $\{A \rightarrow B, B \rightarrow C, B \rightarrow D, C \rightarrow F, D \rightarrow F, D \rightarrow G\}$ y las siguientes probabilidades:

$$\begin{aligned}P(+a) &= 0'1, \\P(+b|+a) &= 0'8, \quad P(+b|\neg a) = 0'25, \\P(+c|+b) &= 0'7, \quad P(+c|\neg b) = 0'35, \\P(+d|+b) &= 0'6, \quad P(+d|\neg b) = 0'1, \\P(+f|+c, +d) &= 0'8, \quad P(+f|+c, \neg d) = 0'6, \quad P(+f|\neg c, +d) = 0'5, \quad P(+f|\neg c, \neg d) = 0, \\P(+g|+d) &= 0'4, \quad P(+g|\neg d) = 0'1.\end{aligned}$$

1. Calcule la probabilidad $P(b|\neg a, +f, \neg g)$ con el método de eliminación de variables, indicando los **valores numéricos** de los potenciales que se calculan en cada paso.
2. Calcule la misma probabilidad utilizando el programa OpenMarkov. Capture la pantalla para esta red en modo inferencia, con los hallazgos introducidos, e inclúyala en sus respuestas.

Ejercicio 2.2.

Indique detalladamente qué pasos seguiría cada uno de los métodos siguientes para calcular la probabilidad anterior. No es necesario que realice cálculos numéricos.

1. Muestreo lógico (llamado “método de aceptación-rechazo” en [1]).
2. Ponderación por verosimilitud (llamado “método de la función de verosimilitud pesante” en [1]).

Ejercicio 2.3.

Sea la red bayesiana de 12 variables dada por los enlaces $\{A \rightarrow D, A \rightarrow F, B \rightarrow F, B \rightarrow G, C \rightarrow G, C \rightarrow H, D \rightarrow I, F \rightarrow I, F \rightarrow J, G \rightarrow J, G \rightarrow K, H \rightarrow K, I \rightarrow L, J \rightarrow L, J \rightarrow M, K \rightarrow M\}$. Indique detalladamente los pasos necesarios para calcular la probabilidad a posteriori de A dado cierto valor de L (por ejemplo, $+l$) mediante los siguientes métodos. [Se supone que conocemos las tablas de probabilidades condicionadas que definen la red: $P(a)$, $P(d|a)$, etc., pero en el enunciado no damos sus valores numéricos porque sólo queremos que explique cómo se calculan los potenciales.] Si lo desea, en cada método puede eliminar los nodos sumidero antes de iniciar la inferencia.

1. Eliminación de variables.
2. Agrupamiento (indique también cómo se construye el árbol).
3. Inversión de arcos (indique cómo se calculan las nuevas tablas de probabilidad).

Referencias

- [1] E. Castillo, J. M. Gutiérrez y A. S. Hadi. *Sistemas Expertos y Modelos de Redes Probabilísticas*. Academia de Ingeniería, Madrid, 1997.

ChatGPT 3.5 (29-3-2024)

You

Cuáles son las diferencias principales entre el algoritmo K2 y el PC.

ChatGPT

El algoritmo K2 y el algoritmo PC son dos métodos ampliamente utilizados en el contexto del aprendizaje automático y la inferencia de estructuras de redes bayesianas. Aquí están las principales diferencias entre ellos:

- | | |
|--------------------------------------|---|
| 1. Metodología: | <ul style="list-style-type: none">El algoritmo K2 es un algoritmo de búsqueda exhaustiva que explora todas las posibles estructuras de redes bayesianas ordenadas de manera jerárquica.El algoritmo PC (Constraint-based) se basa en la búsqueda de independencias condicionales en los datos para eliminar conexiones que no sean posibles dadas ciertas restricciones. |
| 2. Complejidad Computacional: | <ul style="list-style-type: none">K2 suele ser más costoso computacionalmente, ya que explora todas las estructuras posibles de manera sistemática.PC tiende a ser más eficiente en términos de tiempo de cómputo ya que se basa en la detección de independencias condicionales en los datos para restringir las posibles estructuras. |
| 3. Uso de Restricciones: | <ul style="list-style-type: none">K2 no se basa en restricciones explícitas, sino que construye la red bayesiana de acuerdo con un orden específico de las variables.PC utiliza restricciones basadas en independencias condicionales encontradas en los datos para eliminar arcos que no son posibles dadas estas restricciones. |
| 4. Resultados: | <ul style="list-style-type: none">K2 puede encontrar una red bayesiana óptima dentro del espacio de búsqueda definido, pero esto depende del orden de las variables y puede no ser óptimo en todos los casos.PC no garantiza una red bayesiana óptima pero puede proporcionar una estructura más precisa al basarse en relaciones de independencia condicional en los datos. |
| 5. Flexibilidad y Ajuste: | <ul style="list-style-type: none">K2 es menos flexible en términos de adaptarse a diferentes tipos de datos y estructuras.PC es más flexible ya que se basa en las relaciones de independencia observadas en los datos. |

En resumen, mientras que el algoritmo K2 es más exhaustivo pero menos eficiente, el algoritmo PC es más rápido y flexible, aunque no garantiza la optimización global de la red bayesiana. La elección entre estos algoritmos depende de las características específicas de los datos y los objetivos del modelado.

You

¿Qué más diferencias hay?

ChatGPT

Además de las diferencias mencionadas anteriormente, hay otros aspectos que distinguen el algoritmo K2 del algoritmo PC en el contexto del aprendizaje de redes bayesianas:

6. Dependencia del Dominio de los Datos:

- K2 generalmente no necesita información previa sobre las dependencias entre variables en el dominio de los datos. Simplemente sigue un proceso sistemático para buscar la mejor estructura de red bayesiana.
- PC requiere que se realice un análisis de independencia condicional en los datos para identificar posibles relaciones entre variables. Esto implica que el usuario necesita tener cierto conocimiento del dominio y de las relaciones esperadas entre las variables.

7. Capacidad para Manejar Datos Incompletos o Ruidosos:

- K2 puede tener dificultades para manejar datos incompletos o ruidosos, ya que explora exhaustivamente todas las estructuras posibles y puede ser sensible a la calidad de los datos.
- PC puede ser más robusto frente a datos incompletos o ruidosos, ya que se basa en la detección de independencias condicionales en los datos disponibles.

8. Escalabilidad:

- K2 puede volverse computacionalmente costoso a medida que aumenta el número de variables en el conjunto de datos, ya que debe explorar todas las posibles estructuras de red.
- PC puede ser más escalable, especialmente en conjuntos de datos grandes, ya que su tiempo de ejecución no depende tanto del número de variables, sino de la complejidad de las relaciones de independencia condicional en los datos.

9. Sensibilidad a la Elección de Parámetros:

- K2 generalmente no requiere ajuste de parámetros y sigue un proceso determinista para encontrar la estructura óptima de red.
- PC puede ser sensible a la elección de parámetros, como el nivel de significancia estadística utilizado para determinar las independencias condicionales en los datos.

Estas diferencias adicionales resaltan las fortalezas y limitaciones específicas de cada algoritmo y destacan la importancia de seleccionar el método más adecuado según las características del conjunto de datos y los objetivos del análisis.

ChatGPT Plus

(3-1-2025)

Pregunta:

Cuáles son las diferencias principales entre el algoritmo K2 y el PC.

ChatGPT Plus

Los algoritmos K2 y PC son métodos utilizados en el aprendizaje de redes bayesianas, pero tienen enfoques y características fundamentalmente diferentes. A continuación, se destacan las principales diferencias:

1. Enfoque General

- **K2:**
 - Es un **algoritmo basado en búsqueda directa y puntuación**.
 - Construye la estructura de la red bayesiana de manera **greedy** (codicioso), seleccionando iterativamente los mejores padres para cada nodo basándose en una función de puntuación.
 - Asume que los nodos están preordenados, lo que significa que el usuario debe proporcionar un orden causal aproximado de las variables.
- **PC (Peter-Clark):**
 - Es un **algoritmo basado en restricciones**.
 - Construye la estructura de la red utilizando pruebas de independencia condicional para inferir las relaciones entre los nodos.
 - No requiere un orden causal previo, sino que descubre la estructura a partir de los datos.

2. Dependencia de Pruebas de Independencia

- **K2:**
 - No utiliza pruebas de independencia condicional. Se basa exclusivamente en la maximización de una métrica de puntuación (como BIC o AIC) para determinar la estructura.
- **PC:**
 - Utiliza pruebas de independencia condicional para identificar las relaciones entre nodos y determinar las conexiones.
 - Requiere un conjunto de datos suficientemente grande para que las pruebas de independencia sean confiables.

3. Complejidad Computacional

- **K2:**
 - Generalmente más eficiente en términos computacionales porque limita la búsqueda de padres a un conjunto predefinido y utiliza una búsqueda local.

- La complejidad depende del número máximo de padres permitido para cada nodo.
 - **PC:**
 - Más costoso computacionalmente debido a la necesidad de realizar múltiples pruebas de independencia condicional para todas las combinaciones posibles de nodos y subconjuntos condicionales.
 - Su complejidad crece exponencialmente con el número de variables y la profundidad condicional evaluada.
-

4. Robustez ante Datos

- **K2:**
 - Funciona bien con datos completos y se adapta a métricas de puntuación que permiten incorporar información previa.
 - Puede ser más sensible al preorden de las variables, lo que afecta su rendimiento si el orden no refleja las verdaderas relaciones causales.
 - **PC:**
 - Es más robusto en la identificación de independencias y dependencias verdaderas, pero su desempeño puede degradarse si las pruebas de independencia no son confiables (por ejemplo, con tamaños de muestra pequeños o ruido en los datos).
-

5. Estructura Resultante

- **K2:**
 - Genera una red bayesiana completamente orientada (DAG - Directed Acyclic Graph).
 - **PC:**
 - Produce inicialmente una estructura parcialmente orientada (CPDAG - Completed Partially Directed Acyclic Graph), que puede ser refinada para convertirse en un DAG en una etapa posterior.
-

6. Aplicaciones

- **K2:**
 - Es útil cuando se tiene conocimiento previo sobre el orden causal de las variables.
 - Adecuado para problemas donde el énfasis está en la puntuación de modelos más que en pruebas de independencia.
 - **PC:**
 - Es útil cuando no se tiene un conocimiento previo del orden causal, y se busca inferir la estructura basándose en las dependencias estadísticas de los datos.
-

En resumen, K2 es más adecuado para situaciones donde se tiene un preorden y se busca una solución eficiente, mientras que PC es preferido cuando se quiere descubrir relaciones causales

directamente de los datos mediante pruebas de independencia. La elección entre ambos depende del contexto y los recursos disponibles.

Tú dijiste:

¿Qué versión de ChatGPT eres?

ChatGPT Plus

Soy ChatGPT con capacidades avanzadas, actualizado a una versión posterior a GPT-4. Puedo realizar tareas como generar imágenes, ejecutar código, recordar información durante una conversación, y acceder a información actual a través de búsquedas en la web. Si necesitas algo específico o más detalles sobre mis habilidades, ¡puedes preguntar! 😊