Autor: Nikita Polyanskiy

Sesión 28 de octubre 2021

Tema 5: Parte 2

Red Bayesiana

Un grafo acíclico dirigido para representar dependencias entre variables y mostrar una descripción escueta de cualquier distribución de probabilidad conjunta completa.

Está formada por:

Un conjunto de variables aleatorias que forman los nodos de la red. Cada nodo X tendrá adjunta una distribución P(X | Padres(X))

Un conjunto de enlaces que determinan la influencia (dependencia) entre nodos. Si X se conecta con Y se dice que X influencia a Y.

Y -> Padre de X

Sabemos que:

$$P(A|B)\,P(B) = P(A,B)$$
$$P(B|A)\,P(A) = P(B,A) = P(A,B)$$

Regla de Bayes

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} = \alpha \cdot P(B \mid A)P(A)$$

α = 1 / P(B)

Cobertura de Markov:

Un nodo A es condicionalmente dependiente (ámbito de influencia) de:

Sus padres,

Sus hijos,

Los padres de sus hijos,

Del resto de nodos es independiente

Con la distribución conjunta (con redes bayesianas) podremos contestar a cualquier pregunta relativa de la red.

Ej.:

A B
O-O-O-O-O
Podremos obtener P(A|B)

La inferencia exacta general funciona para todas las redes bayesianas, pero supone mucha complejidad.

Regla de inferencia general:

(Donde B son las variables buscadas, C las conocidas y D las desconocidas)

$$P(B \mid C) = \alpha \cdot \sum_{D} P(B, D, C)$$

Inferencia aproximada

Las redes con conexión múltiples no se pueden tratar utilizando la inferencia exacta, por lo que se utiliza la inferencia aproximada.

Se puede utilizar varios algoritmos:

Muestreo directo:

Obtiene una muestra aleatoria para cada variable, y se almacena en el vector de sucesos.

Por ejemplo:

P(Terremoto) = <0.001, 0.999> -> Suponemos falso (0.999)

P(Robo) = <0.002, 0.998> -> suponemos falso (0.998)

P(Alarma | T=Falso, R=Falso) = <0.001 , 0.999> -> suponemos cierto(0.001)

P(Juan | A = Cierto) = <0.9, 0.05> -> suponemos cierto (0.9)

P(María | A=Cierto) = <0.7, 0.01> -> suponemos falso (0.01)

X = <falso,falso,cierto,cierto,falso> (vector de sucesos)

Luego para obtener la probabilidad de un suceso contamos las veces que se cumpla el patrón pedido en el vector de sucesos.

P(R|J,M) -> patrón X=<?,Cierto,?, Cierto, Cierto>

C = núm. de veces que se cumple el patrón.

Devuelve C / numTotalDeMuestras

Muestreo por rechazo:

Se apoya en muestreos directos, extrayendo solamente los que coinciden con nuestra evidencia.

Por ejemplo:

P(Lluvia | Aspersor = cierto)

Extraemos 100 muestras, de las cuales 73 tienen Aspersor=falso.

Nos quedamos con 27 donde Aspersor = cierto.

Y de estas 27, en 8 -> lluvia = cierto, en 19 -> lluvia = falso

P(Lluevia | Aspersor = cierto) = <8/27, 19/27> = <0.292, 0.704>

Muestreo de Gibbs:
Se apoya en muestreos por la cobertura de Markov.

Enlace relacionado: https://www.xatakaciencia.com/computacion/asi-se-usan-redes-bayesianas-para-hacer-funcionar-sistemas-expertos-ia

Este articulo explica la definición de los sistemas expertos y las redes bayesianas, y su relación con la IA, explica que un tipo de sistema experto ampliamente usado hoy en día es el basado en una red bayesiana.

Este modelo resulta idóneo para la clasificación, la predicción o el diagnostico, ya que consiste en un grafico que representa un conjunto de variables conocidas y las relaciones de dependencia entre ellas, con el fin de estimar la probabilidad de las variables no conocidas.