Inteligencia Computacional

Razonamiento probabilístico y Redes Bayesianas

Probabilidad y Teorema de Bayes

Blanca A. Vargas Govea - vargasgovea@itesm.mx - Oct 9, 2012

Avisos:

Exámenes

- Fecha de examen parcial 2
  - Martes Octubre 23
- Fecha de examen final
  - Jueves Noviembre 22

Restan 13 sesiones

Proyecto

- Entregar avance proyecto: Martes Octubre 16
- Presentación avance Viernes Nov 9
- Entrega final proyecto Martes Nov 20

Sobre la tarea de AGs

- Esquemas: base sobre la que funcionan los AGs.
- En la implementación de un AG no se programan explícitamente los esquemas.
- Resultados obtenidos en las tareas: diversos debido a los operadores utilizados.

Objetivo

Probabilidad y Teorema de Bayes

Probabilidad de un evento

- Es una especificación del mundo sobre el cual el agente tiene incertidumbre.
- Es la asignación de valores a las variables del mundo.
- Espacio muestral: es el conjunto de todos los resultados posibles.

Probabilidad p de un evento a

$$p(a) = \frac{s}{n}$$

donde s es el número de veces que ocurre a y n es el total de ocurrencias posibles.

Ejemplo: al lanzar un dado, un número par tiene 3 posibles maneras de aparecer de un total de 6

p = 3/6 = 1/2

Operaciones entre conjuntos

Se pueden combinar eventos para formar nuevos.

- $A \cup B$ ocurre si A ocurre o B ocurre o ambos ocurren
- $A \cap B$ ocurre si A ocurre y B ocurre
- A^c complemento de A, ocurre si A no ocurre

A y B son mutuamente excluyentes si son disjuntos

- $A \cap B = \emptyset$ no pueden ocurrir simultáneamente

Operaciones entre conjuntos

Lanzar un dado y observar el número. El espacio muestral son los 6 números posibles:

$S = \{1,2,3,4,5,6\}$

Sea A el evento de salir un número par, B el evento de salir un número impar y C el de salir un número primo:

$A = \{2,4,6\}$, $B = \{1,3,5\}$, $C = \{2,3,5\}$

$A \cup C = \{2,3,4,5,6\}$ par o primo

$B \cap C = \{3,5\}$ primo impar

$C^c = \{1,4,6\}$ no salir un número primo

A y B son mutuamente excluyentes: $A \cap B = \emptyset$

Axiomas

Las probabilidades de los eventos deben satisfacer los siguientes tres axiomas:

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. $P(A \cup B) = P(A) + P(B)$

   $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

   La intersección se cuenta dos veces por eso se resta.

También se les conoce como axiomas de Kolmogorov en honor al matemático ruso Andrei Kolmogorov quien estableció las bases de la teoría de la probabilidad.

Ejemplo

- Una clase contiene 10 hombres y 20 mujeres, de los cuales la mitad de los hombres y la mitad de las mujeres tienen los ojos castaños. Encontrar la probabilidad p de que una persona escogida al azar sea un hombre o tenga los ojos castaños.
  - Sea A = {la persona es un hombre} y
  - B = {la persona tiene ojos castaños}
  - Buscamos p(AUB)
  - $P(A) = 10/30=1/3$, $P(B)=15/30=1/2$ $P(A\cap B)=5/30=1/6$
  - $P(A\cup B) = P(A)+P(B)-P(A\cap B)=1/3+1/2-1/6=0.67$

Inferencia probabilista usando distribuciones conjuntas completas

- Método simple para la inferencia probabilista.
- Sirve para el cálculo de probabilidades posteriores a partir de la evidencia observada.
- Ejemplo: variables booleanas enganche, caries, dolor (de muela).

	dolor		~dolor	
	enganche	~enganche	enganche	~enganche
caries	0.108	0.012	0.072	0.008
~caries	0.016	0.064	0.144	0.576

Todo suma 1

Inferencia probabilista usando distribuciones conjuntas completas

	dolor		~dolor	
	enganche	~enganche	enganche	~enganche
caries	0.108	0.012	0.072	0.008
~caries	0.016	0.064	0.144	0.576

- **Se puede calcular la probabilidad de cualquier proposición.**

Por el axioma 3:

Se identifican los eventos en los que la proposición es verdadera y se suman.

P(caries U dolor-m) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28

Inferencia probabilista usando distribuciones conjuntas completas

	dolor		~dolor	
	enganche	~enganche	enganche	~enganche
caries	0.108	0.012	0.072	0.008
~caries	0.016	0.064	0.144	0.576

- Se puede calcular la distribución sobre cualquier variable o conjunto de variables (probabilidad marginal)

P(caries) = 0.108 + 0.012 + 0.072 + 0.008 = 0.2

¿Es recomendable el método de inferencia?

Probabilidad condicional

- Una vez que se obtiene evidencia, se utilizan las probabilidades condicionales o posteriores.
- Se usa la notación P(a|b), sonde a y b son proposiciones cualquiera.
- Se lee:
  - La probabilidad de a dado que todo lo que conocemos es b
- P(a) es un caso en el que la probabilidad se condiciona a ninguna evidencia.

Probabilidad condicional

- Puede definirse en términos de probabilidades incondicionales
  - $p(a|b) = \frac{p(a \cap b)}{p(b)}$ siempre que p(b) > 0
  - Se puede escribir como $p(a \cap b) = p(a|b)p(b)$
  - … y también como $p(a \cap b) = p(b|a)p(a)$
- Conocida como regla del producto
- Para variables multivaluadas se puede usar la notación
  - $p(X, Y) = p(X|Y)p(Y)$ que combina lo siguiente:
  - $p(X = x_1 \ \cap Y = y_1) = p(X = x_1|Y = y_1)p(Y = y_1)$
  - $p(X = x_1 \ \cap Y = y_2) = p(X = x_1|Y = y_2)p(Y = y_2)$
  - …

Probabilidad condicional

	dolor		~dolor	
	enganche	~enganche	enganche	~enganche
caries	0.108	0.012	0.072	0.008
~caries	0.016	0.064	0.144	0.576

- Aplicando la regla del producto a la tabla del problema dental.
- Podemos calcular la probabilidad de que exista una caries dado que el paciente tiene dolor de muela:
  - $p(caries|dolor) = \frac{p(caries \cap dolor)}{p(dolor)} = \frac{0.108+0.012}{0.108+0.012+0.016+0.064} = 0.6$

Probabilidad condicional

	dolor		~dolor	
	enganche	~enganche	enganche	~enganche
caries	0.108	0.012	0.072	0.008
~caries	0.016	0.064	0.144	0.576

- Podemos calcular la probabilidad de que no hay caries dado que el paciente tiene dolor de muela:
  - $p(\sim caries|dolor) = \frac{p(\sim caries \cap dolor)}{p(dolor)} = \frac{0.016+0.064}{0.108+0.012+0.016+0.064} = 0.4$

El diagnóstico más probable es…

La forma tabular no es práctica. ¿Construirías tu sistema probabilista con una tabla?

Independencia

- Una proposición b es independiente de una proposición a si la probabilidad de que b ocurra no está influenciada por el hecho de que a haya sucedido.
  - $p(Clima = soleado \mid dolor) = p(Clima = soleado)$
- El clima es independiente de los problemas dentales.
- $p(a|b) = p(a)$
- $p(b|a) = p(b)$
- $p(a \cap b) = p(a)\, p(b)$
- $p(X|Y) = p(X)$
- $p(Y|X) = p(Y)$
- $p(X,Y) = p(X)\, p(Y)$

Se basa en el conocimiento del dominio. Puede reducir la información necesaria para especificar la distribución conjunta completa. Aún asi, puede ser mucha.

Teorema de Bayes

- De la regla del producto: $p(a \cap b) = p(a|b)p(b)$
  - $p(a \cap b) = p(a|b)p(b)$
  - $p(a \cap b) = p(b|a)p(a)$ igualando :
  - $p(b|a)p(a) = p(a|b)p(b)$
  - $p(b|a) = \frac{p(a|b)p(b)}{p(a)}$ Teorema de Bayes

Constituye la base de los sistemas modernos de Inteligencia Artificial de inferencia estadística.

Aplicación: caso simple

- Aparentemente no se ve útil.
- Requiere una probabilidad condicional y dos incondicionales.
- En la práctica es útil porque hay muchos casos en los que se tiene la estimación de probabilidades para esas probabilidades.

Aplicación: caso simple

- El doctor sabe que la meningitis origina rigidez en el cuello en un 50% del tiempo.
- El doctor también conoce algunos hechos incondicionales:
  - La probabilidad de que un paciente tenga meningitis es de 1/50000.
  - La probabilidad de que un paciente tenga rigidez en el cuello es de 1/20.
- Sea s la proposición de que el paciente tenga el cuello rígido y m la que el paciente tenga meningitis
  - $p(s|m) = 0.5$
  - $p(m) = \frac{1}{50000}$
  - $p(s) = \frac{1}{20}$
  - $p(m|s) = \frac{p(s|m)p(m)}{p(s)} = \frac{0.5*1/50000}{1/20} = 0.0002$
- Esperamos solamente 1/50000 pacientes con rigidez en el cuello que tengan meningitis.

Ejercicios y tarea

- Se encuentran aquí.
- Ejercicios para hoy.
- Tarea para el Viernes 12.

Referencias

- Russell, S., y Norvig, P. (2003). Artificial intelligence: A modern approach (2nd edition ed.). Prentice-Hall, Englewood Cliffs, NJ.