

# Aprendizaje Bayesiano

## Apéndice A

Oscar Javier Prieto Izquierdo

Raúl Casillas Díez

# Redes bayesianas

- El aprendizaje bayesiano basado en el clasificador naive toma la asunción de la independiencia de las variables.
- Esta asunción es demasiado restrictiva.

# Redes bayesianas (II)

- Las redes bayesianas actúan como un punto intermedio entre el clasificador naive y no hacer suposición alguna sobre la independencia de los atributos.
  - Permiten representar suposiciones de independencia para subconjuntos de variables.
  - Nos dan un conocimiento a priori sobre la dependencia de los atributos.

# Redes bayesianas (III)

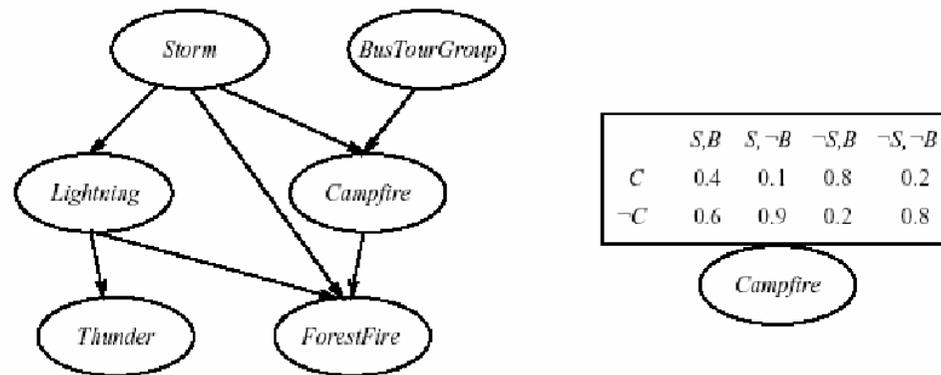
- Si tenemos tres variables aleatorias  $X$ ,  $Y$ ,  $Z$  con valores discretos, decimos que  $X$  es independiente de  $Y$  con respecto a  $Z$  si la distribución de probabilidad de  $X$  es independiente de  $Y$  dado un valor  $Z$ .

$$(\forall x_i, y_j, z_k) P(X = x_i | Y = y_j, Z = z_k) = P(X = x_i | Z = z_k)$$

- Por ejemplo, *Truenos es independiente de Lluvia dado Relámpagos.*

# Representación

- Es un grafo acíclico + tabla de probabilidad condicional.
- El grafo representa las suposiciones de dependencia
  - Una variable es independiente de las que no son predecesoras en el grafo.
- Con cada variables (nodo del grafo) se asocia una tabla con las probabilidades que proporciona la distribución de probabilidades de esa variable dados los valores de sus predecesoras inmediatas.



# Representación (II)

- La probabilidad de una asignación de valores a las variables:

$$P(y_1, \dots, y_n) = \prod_{i=1}^n P(y_i | \text{Padres}(Y_i))$$

donde  $\text{Padres}(Y_i)$  representa a los predecesores inmediatos de  $Y_i$ .

- Las tablas representan las probabilidades  $P(y_i | \text{Padres}(Y_i))$

$P(\text{Campfire} = \text{true} | \text{Storm} = \text{true}, \text{BusTourGroup} = \text{true}) = 0.4$

	$S, B$	$S, \neg B$	$\neg S, B$	$\neg S, \neg B$
$C$	0.4	0.1	0.8	0.2
$\neg C$	0.6	0.9	0.2	0.8

*Campfire*

# Otras cuestiones

- **Inferencia:** Se pueden inferir las distribuciones de probabilidad de unas variables a partir de los valores de las otras
  - Fácil para una sola variable.
- **Aprendizaje:** Sencillo cuando las variables con observables y la estructura es conocida.