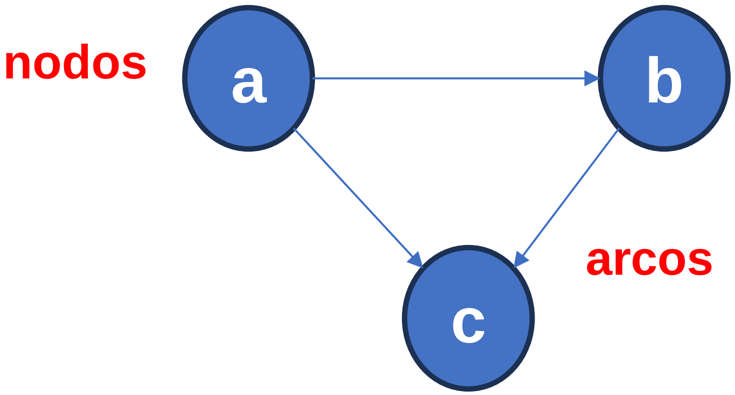
Módulo 4: Aprendizaje Automático

Temario de la Clase 8
27 de septiembre de 2024
Redes Bayesianas

- Modelos gráficos
- Teorema de Bayes
- Redes Bayesianas
- Aplicaciones en Python

Modelos Gráficos

Definición: Los *modelos gráficos* son una representación probabilística que utiliza grafos para describir las relaciones entre variables.



Cada nodo representa una variable aleatoria (o un grupo de variables aleatorias) y los enlaces expresan relaciones probabilísticas entre estas variables.

El gráfico captura entonces la forma en que la distribución conjunta de todas las variables aleatorias puede descomponerse en un producto de factores que dependen cada uno solo de un subconjunto de las variables.

Modelos Gráficos

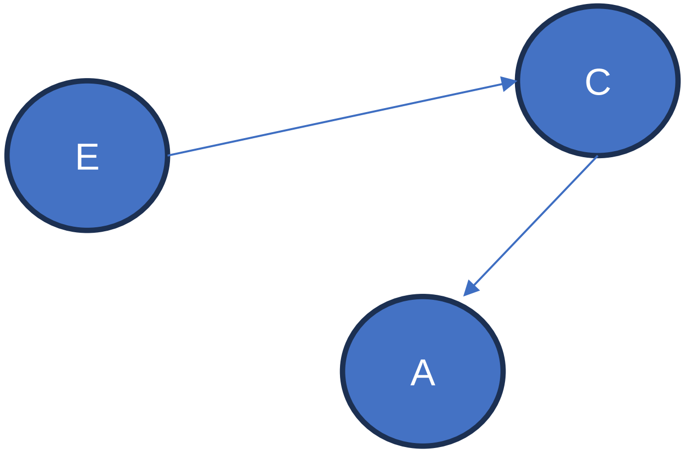
Ejemplo:

Queremos determinar si una persona aprobará un examen. Las variables que vamos a usar son:

E: Estudiar (sí o no).

C: Concentración durante el examen (sí o no).

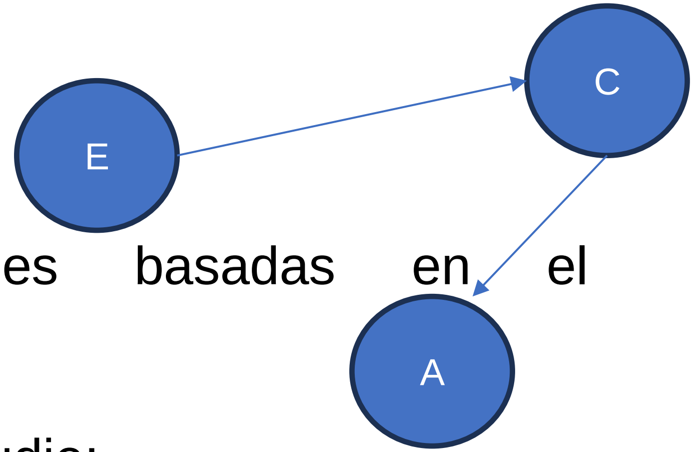
A: Aprobar el examen (sí o no).



La flecha de E hacia C significa que estudiar influye en la probabilidad de estar concentrado durante el examen.

La flecha de C hacia A indica que estar concentrado afecta la probabilidad de aprobar.

Modelos Gráficos



Probabilidades condicionales:

Asignamos probabilidades condicionales basadas en el comportamiento esperado:

- La probabilidad de que una persona estudie:

$$P(E = \text{sí}) = 0.6 \;\; \text{y} \;\; P(E = \text{no}) = 0.4$$

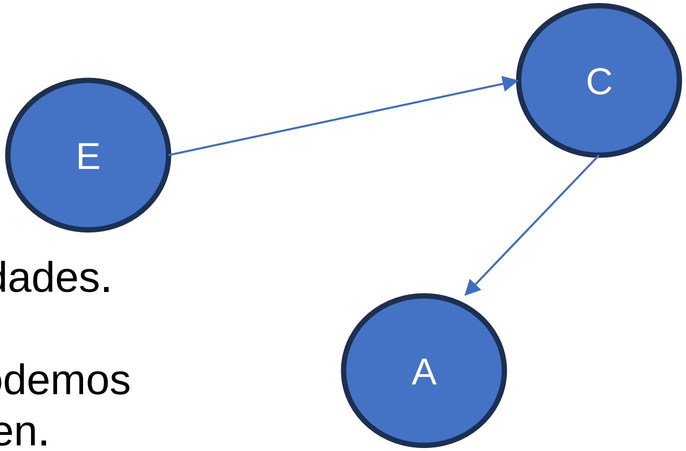
- Si una persona estudia, la probabilidad de que esté concentrada durante el examen:

$$P(C = \text{sí}|E = \text{sí}) = 0.9 \;\; \text{y} \;\; P(C = \text{sí}|E = \text{no}) = 0.4$$

- Si una persona está concentrada, la probabilidad de aprobar el examen:

$$P(A = \text{sí}|C = \text{sí}) = 0.8 \text{ y } P(A = \text{sí}|C = \text{no}) = 0.2$$

Modelos Gráficos



Inferencia :

Podemos usar el modelo para calcular probabilidades.

Por ejemplo, si sabemos que alguien estudió, podemos calcular la probabilidad de que apruebe el examen.

Primero, calculamos la probabilidad de estar concentrado dado que estudió:

$$P(C = \text{sí} | E = \text{sí}) = 0.9 \quad P(C = \text{sí} \; | E = \text{sí}) = 0.9$$

Luego, podemos calcular la probabilidad de aprobar dado que estudió, usando las probabilidades condicionales:

$$P(A = \text{sí} | E = \text{sí}) = P(A = \text{sí} | C = \text{sí}) \cdot P(C = \text{sí} | E = \text{sí}) + P(A = \text{sí} | C = \text{no}) \cdot P(C = \text{no} | E = \text{sí})$$

Sustituyendo los valores:

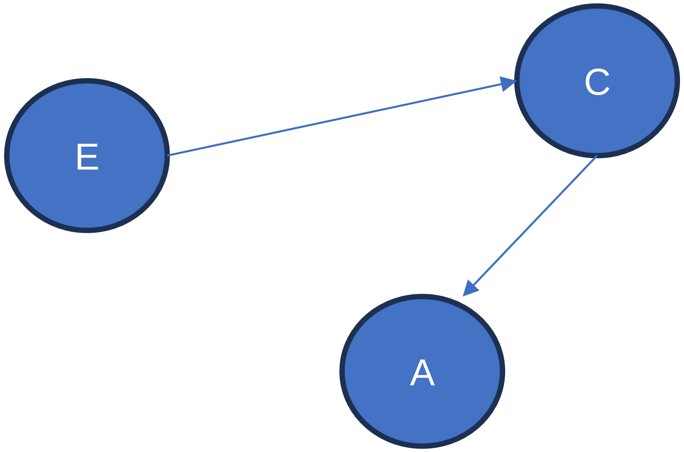
$$P(A = \text{sí} | E = \text{sí}) = (0.8 \cdot 0.9) + (0.2 \cdot 0.1) = 0.72 + 0.02 = 0.74$$

Esto significa que si alguien estudia, hay un 74% de probabilidad de que apruebe el examen.

Modelos Gráficos

Conclusión:

Este ejemplo muestra cómo los modelos gráficos pueden representar dependencias (en este caso, Estudiar → Concentración → Aprobar) y cómo las probabilidades condicionales nos permiten hacer inferencias sobre el éxito en el examen, dado que sabemos si una persona estudió o no.



Teorema de Bayes

Sea $\{A_1, A_2, \ldots; A_i, \ldots, A_n\}$ un conjunto de sucesos mutuamente excluyentes y exhaustivos tales que la probabilidad de cada uno de ellos es distinta de cero [$P(A_i) \neq 0$, i=1, 2, .., n]. Si B es un suceso cualquiera del que se conocen las probabilidades condicionales $P(B|A_i)$ entonces la probabilidad $P(A_i|B)$ viene dada por la expresión:

Thomas Bayes (1702-1761)

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^{n} P(B|A_k)P(A_k)}.$$

- $P(A_i)$ son las probabilidades *a priori*,
- $P(B|A_i)$ es la probabilidad de B en la hipótesis Ai
- $P(A_i|B)$ son las probabilidades *a posteriori*.

Teorema de Bayes

$$P(E = sí|A = sí) = \frac{P(A = sí|E = sí) * P(E = sí)}{P(A = sí)})$$

Donde:

- $P(E = \text{sí}|A = \text{sí})$ es la probabilidad de que haya estudiado dado que aprobó.
- $P(A = \text{sí}|E = \text{sí})$ es la probabilidad de aprobar dado que estudió, que ya calculamos en la inferencia anterior como 0.74.
- $P(E = \text{sí}) = 0.6$ es la probabilidad de que alguien estudie.
- $P(A = \text{sí})$ es la probabilidad total de aprobar el examen.

Teorema de Bayes

Para calcular $P(A$ = sí), necesitamos tener en cuenta ambas situaciones posibles: que la persona haya estudiado o no.

P(A = sí) = P(A = sí|E = sí) · P(E = sí) + P(A = sí| E = no) · P(E = no)

Ya tenemos la información:

- P(A = sí|E = sí) = 0.74
- P(E = sí) = 0.6
- P(A = sí|E = no) = P(A = sí|C = sí) · P(C = sí| E = no) + P(A = sí|C = no) ·P(C = no|E = no)

Para el segundo término:

- P(C = sí| E = no) = 0.4
- P(A = sí|C = sí) = 0.8
- P(A = sí|C = no) = 0.2
- P(C = no|E = no) = 1 – 0.4 = 0.6
- P(A = sí|E = no) = (0.8*0.4) + (0.2*06) = 0.44

Ahora P(A = sí) = (0.74*0.6) + (0.44*0.4) = 0.62

Teorema de Bayes

Aplicamos el Teorema de Bayes

Ahora que tenemos $P(A$=sı´)=0.62, podemos aplicar el teorema de Bayes:

$$P(E = sí \,|A = sí) = \frac{0.74 \;\cdot 0.6}{0.62} = 0{,}716$$

Conclusión: La probabilidad de que alguien haya estudiado dado que aprobó el examen es aproximadamente 0.716, es decir, un 71.6%.

Teorema de Bayes permite realizar inferencias inversas, en este caso, deduciendo si alguien estudió basado en el hecho de que aprobó.

Redes Bayesianas

Son un tipo de modelo gráfico probabilístico que representa un conjunto de variables y sus dependencias condicionales mediante un grafo dirigido acíclico (DAG).

Están directamente relacionadas con el Teorema de Bayes, ya que utilizan probabilidades condicionales para hacer inferencias sobre un sistema de variables interconectadas.

Concepto de Redes Bayesianas:

Nodos: Cada nodo del grafo representa una variable aleatoria que puede ser discreta o continua (como “Estudiar” o “Aprobar”).

Arcos/Flechas: Las flechas entre los nodos representan dependencias condicionales entre las variables. Si hay una flecha de una variable A a una variable B, significa que A influye en la probabilidad de que ocurra B.

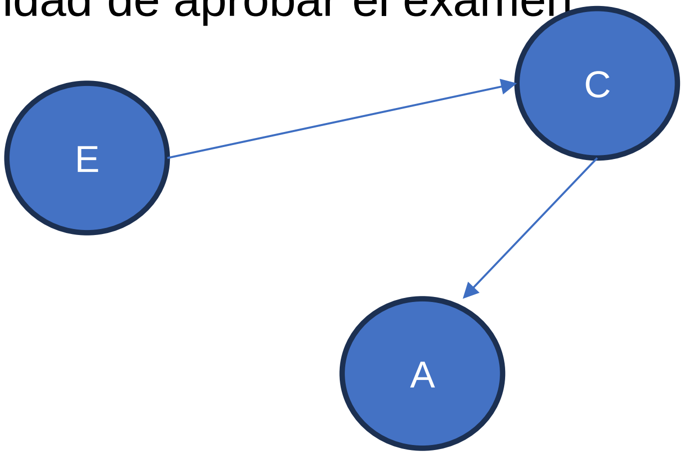
Distribuciones de Probabilidad Condicional: Cada nodo está asociado a una distribución de probabilidad condicional dada la variable de la que depende.

Redes Bayesianas

Ejemplo:
Relación con el problema de Estudiar y Aprobar un Examen.

Tomemos el mismo ejemplo que discutimos antes: Tenemos tres variables:
E (Estudiar); C (Concentración); A (Aprobar el examen).

Estas variables están conectadas mediante flechas que indican relaciones de dependencia: E → C: Estudiar afecta la probabilidad de estar concentrado.
C → A: Estar concentrado afecta la probabilidad de aprobar el examen.
La red bayesiana para este ejemplo es:



Esto significa que:

La probabilidad de estar concentrado $P(C)$ depende de si se estudió $P(C|S)$.
La probabilidad de aprobar $P(A)$ depende de estar concentrado $P(A|C)$.

Redes Bayesianas

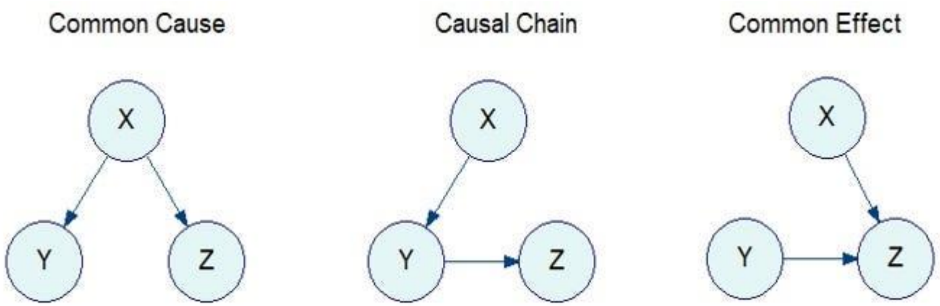
Relación con el Teorema de Bayes:

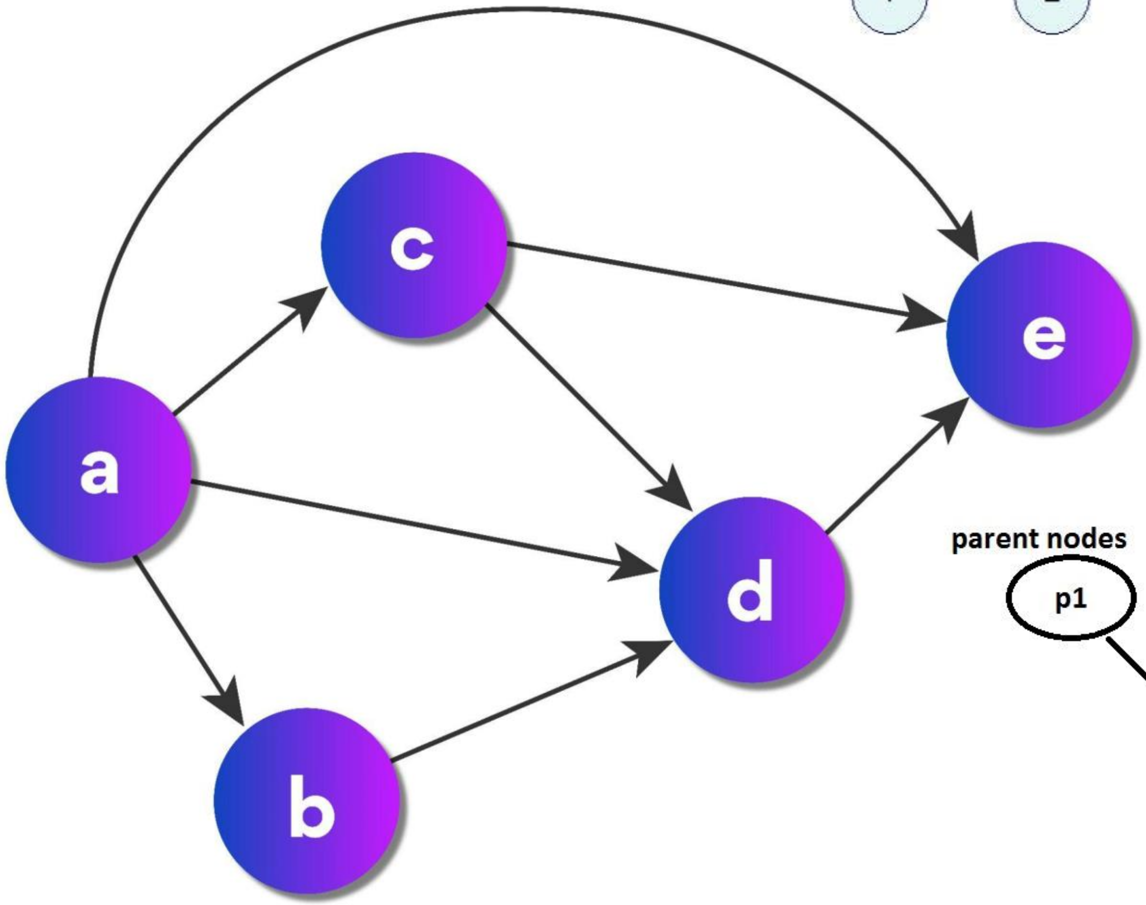
Las redes bayesianas son una aplicación directa del Teorema de Bayes. Cada vez que queremos calcular la probabilidad de un evento dado otro evento (como la probabilidad de aprobar dado que estudió), utilizamos el Teorema de Bayes para realizar esta inferencia.

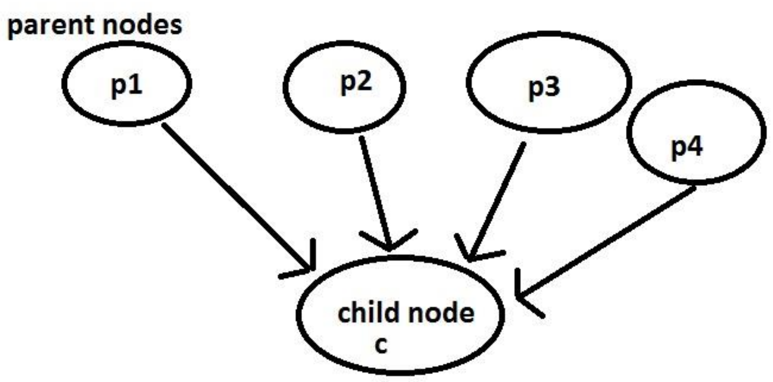
Por ejemplo, en el caso de la red bayesiana sobre estudiar y aprobar:
Si sabemos que alguien aprobó el examen, podemos usar el Teorema de Bayes para calcular la probabilidad de que haya estudiado, utilizando la estructura y las relaciones del grafo.

La red nos permite organizar las relaciones entre variables de manera eficiente y clara. Además, usando el Teorema de Bayes, podemos actualizar nuestras creencias sobre una variable cuando obtenemos nueva información sobre otras.:
Relación con el problema de Estudiar y Aprobar un Examen.

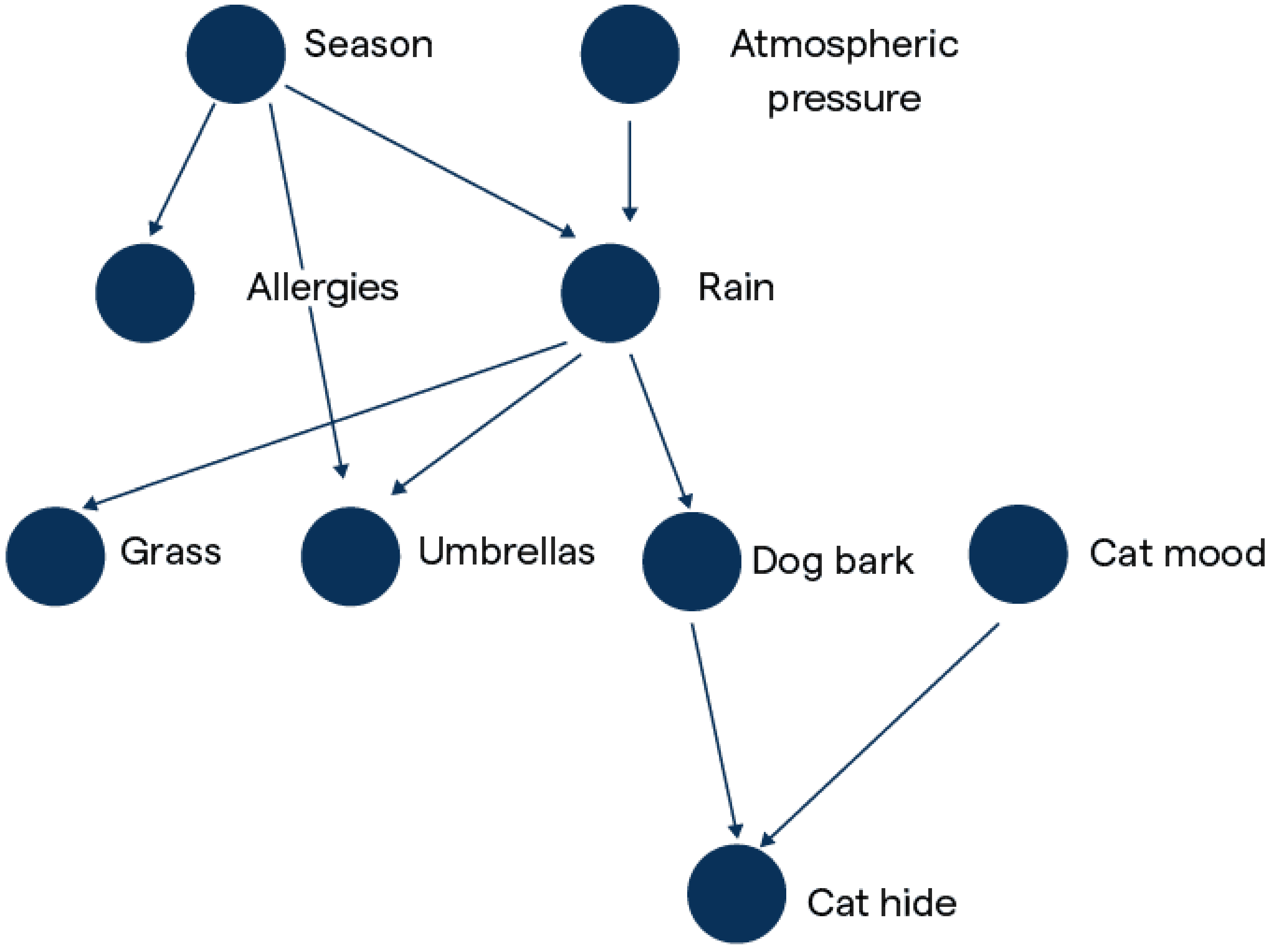
Redes Bayesianas



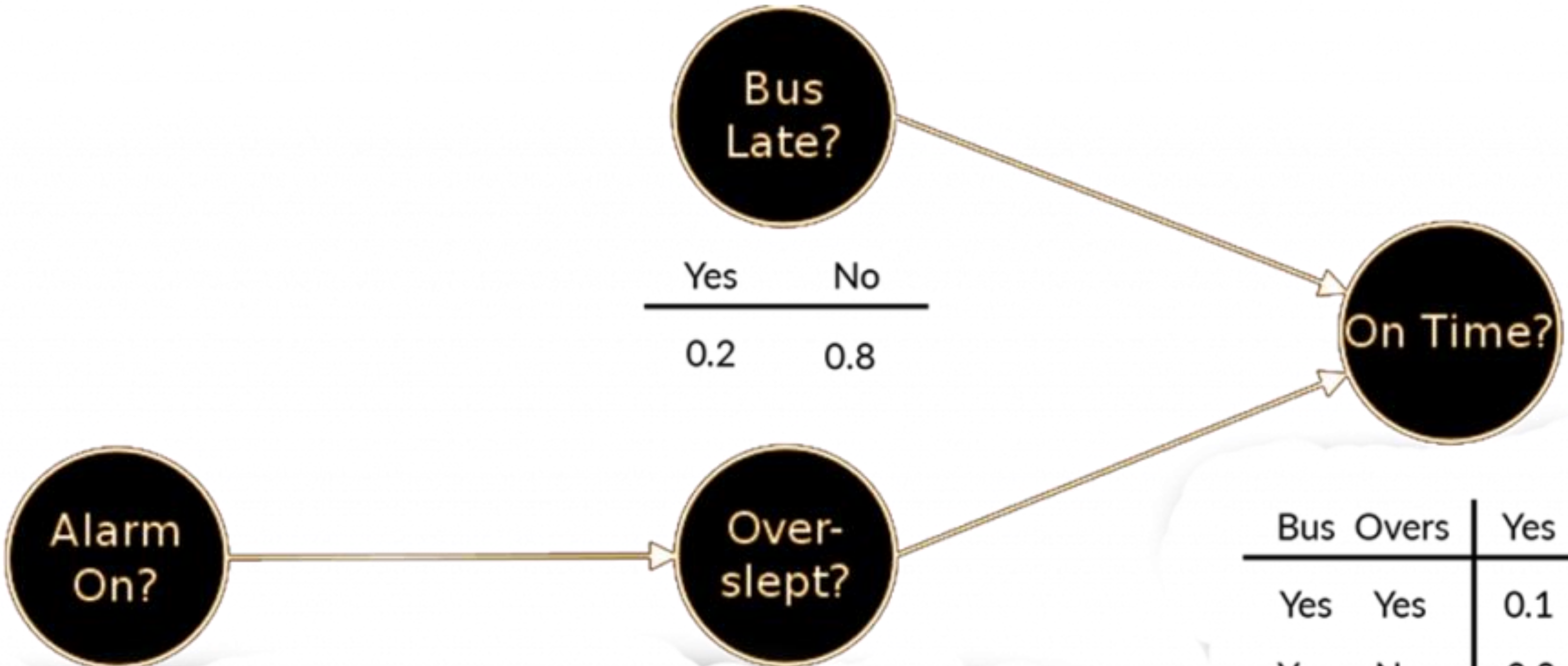




Redes Bayesianas



Redes Bayesianas



Bus Late?

Yes	No
0.2	0.8

Alarm On?

Yes	No
0.9	0.1

Over-slept?

Alarm	Yes	No
Yes	0.1	0.9
No	0.9	0.1

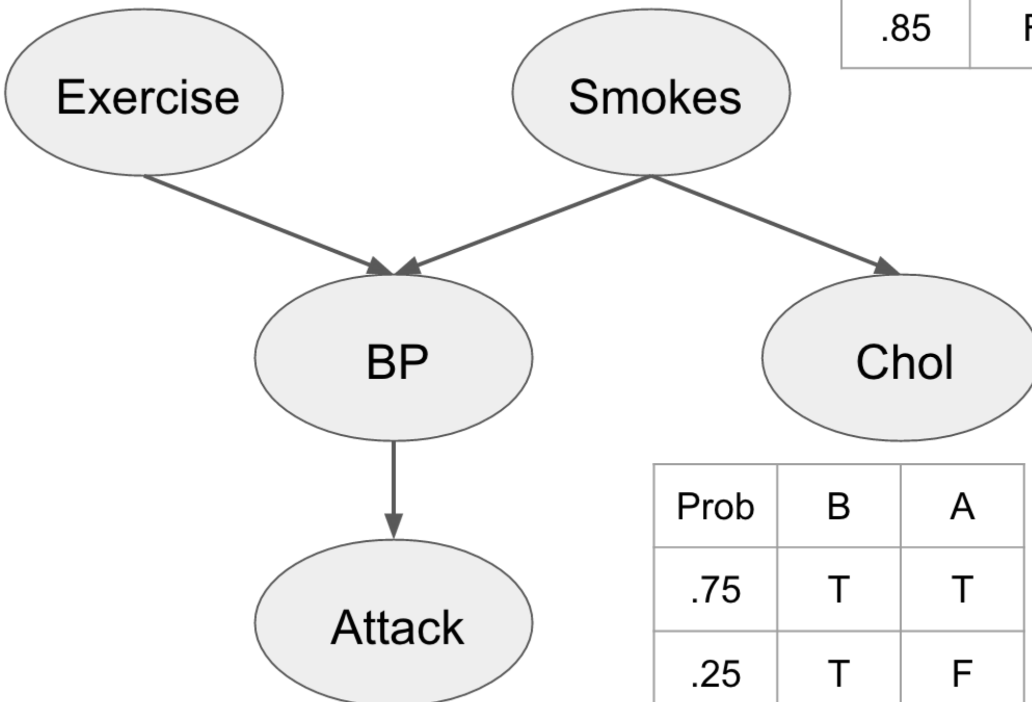
On Time?

Bus	Overs	Yes	No
Yes	Yes	0.1	0.9
Yes	No	0.2	0.8
No	Yes	0.3	0.7
No	No	0.9	0.1

Redes Bayesianas

Prob	E
.4	T
.6	F

Prob	S
.15	T
.85	F



Prob	E	S	B
.45	T	T	T
.55	T	T	F
.05	T	F	T
.95	T	F	F
.95	F	T	T
.05	F	T	F
.55	F	F	T
.45	F	F	F

Prob	S	C
.8	T	T
.2	T	F
.4	F	T
.6	F	F

Prob	B	A
.75	T	T
.25	T	F
.05	F	T
.95	F	F

Redes Bayesianas

¿Cómo las redes bayesianas ayudan en la práctica?

Inferencia: A partir de observaciones parciales (por ejemplo, saber que alguien aprobó), podemos inferir probabilidades sobre otras variables (como si estudió o si estuvo concentrado).

Predicción: Podemos predecir el resultado de una variable sabiendo el estado de otras (por ejemplo, predecir si una persona aprobará sabiendo si estudió y estuvo concentrado).

Toma de decisiones: Las redes bayesianas se utilizan mucho en áreas como el diagnóstico médico, la detección de fraudes y los sistemas de recomendación..

Las redes bayesianas son modelos gráficos que muestran las dependencias condicionales entre variables, y utilizan el Teorema de Bayes para hacer cálculos probabilísticos e inferencias.
En el ejemplo de estudiar y aprobar, una red bayesiana nos permite ver cómo estudiar influye en la concentración y, a su vez, cómo la concentración afecta la probabilidad de aprobar. Usando el Teorema de Bayes, podemos hacer inferencias sobre las probabilidades de estos eventos.