REDES BAYESIANAS

- Hay una hipótesis de independencia, a menudo no son ciertas
- Modela explícitamente las relaciones de independencia de las variables asociadas a los datos.

  Usa relaciones de independencia para hacer inferencias probabilísticas
- Lenguaje intuitivo – usar conocimiento causal en la construcción de los modelos.

Variable aleatoria

Son independientes cuando la distribución de probabilidad conjunta es el producto de las probabilidades marginales.

ESTRUCTURA

- Los nodos representan las variables aleatorias. A priori toman valores discretos
- Los arcos representan una influencia directa probabilística.
- La estructura representa la relación de independencia/dependencia probabilística entre las variables aleatorias

Independencia Condicional

Dados Y, Z y W se dice que Y es condicionalmente independiente de Z dado W si se verifica

que para algún x, y, z : $p(y \mid z, w) = p(y \mid w)$

Para cada variable X_i, la estructura S representa la afirmación de que X_i es independiente a todos los nodos excepto del padre.

$$p(\vec{x}) = p(x_1 \dots x_n) = \underbrace{\prod_i^n p(x_i \mid pa_i^S)}$$

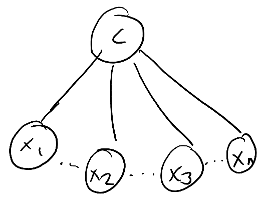
pa_i^S → conjunto de padres de x_i en S.

→ sustituir por la distribución que siga x_i

$$\begin{cases} \prod^n (p_i)^{x_i} (1-p_i)^{1-x_i} \longrightarrow x \in B(p_i) \\ \prod^n (2\pi\sigma_i)^{-1/2} e^{\left(\frac{(x_i-\mu_i)^2}{\sigma_i^2}\right)} \longrightarrow x \in N(\mu_i, \sigma_i^2) \end{cases}$$

Naive Bayes

Se prueba la pertenencia del patrón a cada una de las clases y se escoge la de mayor probabilidad:



Discreta: $c^* = \arg\max P(C=c) \prod^n P(X_i = x_i \mid C = c)$

Continua y normal: $c^* = \arg\max P(C=c) \prod^n \frac{1}{\sqrt{\pi \sigma_i^c}} \cdot e^{-\frac{1}{2}\left(\frac{x_i - \mu_i^c}{\sigma_i^c}\right)^2}$

Información mutua entre X e Y

$$I(X,Y) = \sum_{i}^{r_X} \sum_{j}^{r_Y} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x_i)\,p(y_j)} \xrightarrow{\text{condicionada a } C} I(X,Y|C) = \sum_c p(C=c)\, I(X,Y|C=c) =$$

$$= \sum_{i}^{r_X} \sum_{j}^{r_Y} \sum_{k}^{r_C} p(x_i, y_j, c_k) \log \frac{p(x_i, y_j | c_k)}{p(x_i|c_k)\,p(y_j|c_k)}$$

reducción de la incertidumbre de una de las variables cuando se conoce la otra

Cantidad de información mutua y la verosimilitud