Teoria dell'Informazione

Simone Alessandro Casciaro

29 Ottobre 2024

Lezione 9: Derivati dell'Entropia II

Proseguio Lezione Precedente

Nella Lezione 8, abbiamo detto che se abbiamo tre variabili aleatorie X, Y, Z tali che X, Z sono indipendenti se condizionate a Y, allora

$$I(X, Y) \geq I(X, Z)$$

Tuttavia, ci chiediamo cosa succede nel caso in cui Z non sia indipendente da X.

> Esempio
> Supponiamo X e Y indipendenti, con $Z = X + Y$
> $$X = \begin{cases} 0 \\ 1 \end{cases} \qquad Y = \begin{cases} 0 \\ 1 \end{cases} \qquad Z = \begin{cases} 0 \\ 1 \\ 2 \end{cases}$$
>
> $I(X, Y) = H(X) - H(X|Y) = H(X) - H(X) = 0$ perché X e Y sono indipendenti

$$
\begin{aligned}
I(X,Y|Z) &= H(X|Z) - H(X|Y,Z) \\
&= H(X|Z) && \text{perché } X \text{ è dipendente da } Y, Z \\
&= \sum_{z \in \mathbb{Z}} p(Z=z) H(X|Z=z) \\
&= \mathbb{P}(Z=0)H(X|Z=0) + \mathbb{P}(Z=1)H(X|Z=1) + \mathbb{P}(Z=2)H(X|Z=2) \\
&= \mathbb{P}(Z=1)H(X|Z=1) \\
&= \frac{1}{2} * 1 \\
&= \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
I(X,Z) &= H(X) - H(X|Z) \\
&= H(X) - I(X,Y|Z) \\
&= 1 - \frac{1}{2} \\
&= \frac{1}{2}
\end{aligned}
$$

Quindi abbiamo dimostrato che

$$I(X,Y) \ngeq I(X,Z)$$

Esercizio: Tema d'Esame 1 Aprile 2003

Si supponga di avere lo schema $\mathbb{S} - C - \mathbb{R}$ che identificano rispettivamente la Sorgente, il Canale e Ricevente.
$\mathbb{S} = \{s_1, s_2, s_3, s_4\}$
$\mathbb{P} = \{0.2, 0.3, 0.1, 0.4\}$
Il canale viene rappresentato da una matrice stocastica

$$M = \begin{bmatrix} 0.2 & 0.2 & 0.3 & 0.2 & 0.1 \\ 0.2 & 0.5 & 0.1 & 0.1 & 0.1 \\ 0.6 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.3 & 0.1 & 0.1 & 0.1 & 0.4 \end{bmatrix}$$

dove $i = \{1, \ldots, 4\}$ è l'indice delle righe di M e $j = \{1, \ldots, 5\}$ è l'indice delle colonne.

Calcolare $H_2(R|S)$

$$H_2(R|S) = \sum_{i=1}^{4} p(x_i) \sum_{j=1}^{5} p(y_j|x_i) \log_2 \frac{1}{p(y_j|x_i)}$$

Ogni entrata della matrice M_{ij} rappresenta tutti i vari possibili $p(y_j|x_i)$, mentre i $p(x_i)$ sono dati dalla distribuzione di probabilità $\mathbb{P}$

Possiamo dunque sostituire tutti i valori.

$$\begin{aligned} H_2(R|S) = 0.2 &\left(0.2 \log_2 \frac{1}{0.2} + 0.2 \log_2 \frac{1}{0.2} + 0.3 \log_2 \frac{1}{0.3} + 0.2 \log_2 \frac{1}{0.2} + 0.1 \log_2 \frac{1}{0.1}\right) \\ + 0.3 &\left(0.2 \log_2 \frac{1}{0.2} + 0.5 \log_2 \frac{1}{0.5} + 0.1 \log_2 \frac{1}{0.1} + 0.1 \log_2 \frac{1}{0.1} + 0.1 \log_2 \frac{1}{0.1}\right) \\ + 0.1 &\left(0.6 \log_2 \frac{1}{0.6} + 0.1 \log_2 \frac{1}{0.1} + 0.1 \log_2 \frac{1}{0.1} + 0.1 \log_2 \frac{1}{0.1} + 0.1 \log_2 \frac{1}{0.1}\right) \\ + 0.4 &\left(0.3 \log_2 \frac{1}{0.3} + 0.1 \log_2 \frac{1}{0.1} + 0.1 \log_2 \frac{1}{0.1} + 0.1 \log_2 \frac{1}{0.1} + 0.4 \log_2 \frac{1}{0.4}\right) \\ = &\boxed{2.033} \end{aligned}$$

Trovare $H_2(S|R)$

Partiamo dalla definizione di $H(S|R) = \sum_{j=1}^{5} \sum_{i=1}^{4} p(y_j) p(x_i|y_j) \log \frac{1}{p(x_i|y_j)}$

Sapendo che $p(y_j)p(x_i|y_j) = p(x_i, y_j) = p(x_i)p(y_j|x_i)$, possiamo sostituire ottenendo

$$H(S|R) = \sum_{j=1}^{5} \sum_{i=1}^{4} p(x_i)p(y_j|x_i) \log \frac{p(y_j)}{p(x_i)p(y_j|x_i)}$$

Grazie alla definizione di probabilità marginale, sappiamo che $p(y_j) = \sum_{i=1}^{4} p(x_i, y_j) = \sum_{i=1}^{4} p(x_i)p(y_j|x_i)$ e quindi sostituiamo

$$H(S|R) = \sum_{j=1}^{5} \sum_{i=1}^{4} p(x_i)p(y_j|x_i) \log \left(\frac{\sum_{k=1}^{4} p(x_k)p(y_j|x_k)}{p(x_i)p(y_j|x_i)} \right)$$

Implementiamo questa formula in Python

```
def H_SR(C: list[list[float]], P: list[float]) -> float:
    if len(C) != len(P):
        raise Exception('Lenght are not correct')
    entropy = 0
    for x in range(len(C)):
        for y in range(len(C[x])):
            p_y = sum([C[xx][y]*P[xx] for xx in range(len(C))])
            val = M[x][y]*P[x]
            entropy += p_xy*math.log(p_y/p_xy,2)
    return entropy
```

e il codice restituisce come valore $H(S|R) = \boxed{1.627}$