

## 4 Aprendizaje de parámetros y grafos



Camilo Palazuelos Calderón

REPRESENTACIÓN DEL CONOCIMIENTO  
Grado en Ingeniería Informática  
Mención en Computación

Curso 2023-2024

# Información útil

Sobre la práctica y su entrega

## ■ Objetivos de la práctica

- Profundizar en la manipulación de grafos en Python
- Implementar el algoritmo de aprendizaje de Chow-Liu
- Calcular el coste temporal del algoritmo codificado

## ■ Laboratorio: 24 de noviembre y 1 de diciembre de 14:30 a 16:30

- La fecha límite de entrega es el 7 de diciembre a las 23:59

L	M	X	J	V
13	14	15	16	17
20	21	22	23	24
27	28	29	30	1
4	5	6	7	8
11	12	13	14	15

## ■ Qué entregar

- Memoria con respuestas a las preguntas formuladas en el guion de la práctica
- Código desarrollado (y material adicional si lo consideráis oportuno)

# Estimación de máxima verosimilitud

El caso del grafo de una RB discreta como arborescencia

## ■ Algoritmo de Chow-Liu

- ① Crea  $\mathcal{K} = (V, E, W)$  tal que  $w_{ij} = \mathbb{I}(X_i, X_j \mid \mathcal{D})$  para todo  $\{v_i, v_j\} \in E$
- ② Busca un árbol de recubrimiento de peso máximo  $\mathcal{G} = (V, E')$  de  $\mathcal{K}$
- ③ Elige  $v \in V$  a partir del que asignar direccionalidad a  $\mathcal{G}$  *hacia fuera*

## Información mutua de dos variables dado $\mathcal{D}$

$$\mathbb{I}(X, Y \mid \mathcal{D}) = \sum_{x,y} p_{\mathcal{D}}(x, y) \log \frac{p_{\mathcal{D}}(x, y)}{p_{\mathcal{D}}(x) \cdot p_{\mathcal{D}}(y)}$$

- $\arg \max_{\mathcal{G}} \log p(\mathcal{D} \mid \mathcal{G}) = \arg \max_{\mathcal{G}} \sum_{(v_i, v_j) \in E(\mathcal{G})} \mathbb{I}(X, Y \mid \mathcal{D})$ 
  - Restringir la estructura de  $\mathcal{G}$  evita un *sobreajuste* en el aprendizaje

# Tareas y preguntas

Qué hacer y a qué dar respuesta en la memoria

- 1 [6 PUNTOS] **Codificación del algoritmo descrito**
  - Para ello, os recomiendo utilizar el módulo de Python NetworkX
- 2 [2 PUNTOS] **Eficacia de vuestra propuesta**
  - Mostrad, con ejemplos variados, que funciona correctamente
- 3 [2 PUNTOS] **Evaluación de su coste temporal empírico**
  - En función de algún parámetro de su coste temporal teórico