

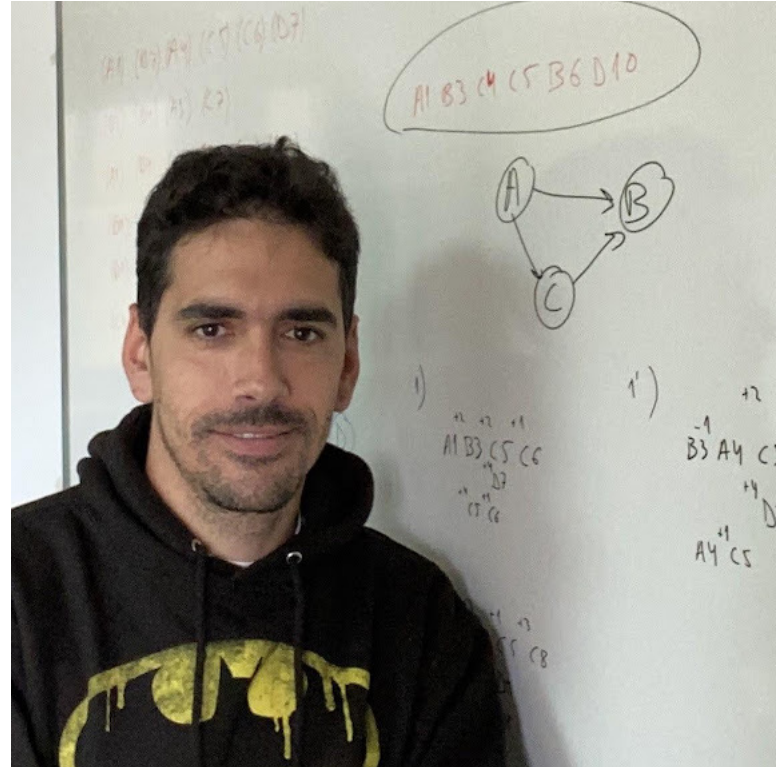
A stylized sunburst graphic in shades of purple and blue, located in the top-left corner of the slide.

Métodos descriptivos supervisados

Máster en Ciencia de Datos

José María Luna

recibió el título de Doctor en Ciencias de la Computación en 2014, por la Universidad de Granada. Su carrera investigadora fue inicialmente subvencionada por el Ministerio de Educación de España bajo el programa FPU (predoctoral) y el programa Juan de la Cierva (postdoctoral). Actualmente es profesor de la Universidad de Córdoba en el departamento de Informática y Análisis Numérico. Dr. Luna ha sido autor de los libros monográficos "Pattern Mining with Evolutionary Algorithms" y "Supervised Descriptive Pattern Mining", ambos publicados por la editorial Springer. Además, ha publicado más de 30 artículos en revistas científicas de alto impacto. Actualmente tiene un total de 2260 citas en Google Scholar y un índice H de 24. Su investigación es llevada a cabo en el grupo de investigación *Knowledge Discovery and Intelligent Systems*, donde investiga temas relativos a computación evolutiva, minería de patrones, reglas de asociación y sus aplicaciones.



A stylized sunburst graphic in shades of purple and blue, located in the top-left corner of the slide. It features a semi-circle on the left with several rays extending outwards to the right.

Métodos descriptivos supervisados (Parte 2)

Emerging Patterns

Patrones emergentes (Emerging patterns)

- Introducción

- La minería de patrones emergentes es una técnica de minería de datos cuya objetivo es el descubrimiento de patrones discriminativos, cuya frecuencia incrementa de manera significativa de un grupo a otro. Cada grupo puede ser un conjunto de datos diferente
- Los patrones extraídos se representan en forma de reglas y representan tendencias emergentes entre conjuntos de datos definidos por una propiedad de interés
- A diferencia de técnicas como las reglas de asociación o el descubrimiento de subgrupos:
 - La minería de patrones emergentes permite contrastar comportamiento diferentes entre grupos, considerando la frecuencia de aparición como principal medida
 - Existen diferentes tipos de patrones emergentes según las características que cumplan

Patrones emergentes (Emerging patterns)

- Definición formal

El problema de minería de patrones emergentes se define como la tarea de descubrir patrones cuya frecuencia de aparición se incrementa considerablemente de un conjunto de datos a otro

$$GR(x) = \begin{cases} 0, & \text{If } Sup_{D_1}(x) = Sup_{D_2}(x) = 0, \\ \infty, & \text{If } Sup_{D_2}(x) \neq 0 \wedge Sup_{D_1}(x) = 0, \\ \frac{Sup_{D_2}(x)}{Sup_{D_1}(x)}, & \text{another case,} \end{cases}$$

Patrones emergentes (Emerging patterns)

- Ejemplos
 - Economía = normal

Sexo	Edad	País	Estudios	Estado civil
Hombre	25	España	Superiores	Soltero
Mujer	45	Alemania	Superiores	Casado
Mujer	62	Italia	Primaria	Casado
Mujer	56	España	Secundaria	Casado

Patrones emergentes (Emerging patterns)

- Ejemplos
 - Economía = rico

Sexo	Edad	País	Estudios	Estado civil
Mujer	37	Suiza	Doctorado	Soltero
Mujer	55	Alemania	Doctorado	Casado
Hombre	71	Suiza	Primaria	Soltero
Hombre	18	Suiza	Secundaria	Soltero

Patrones emergentes (Emerging patterns)

- Ejemplos

- *Pais = Suiza, Estado civil = Soltero*

- *Conjunto de datos Economía = normal → Soporte 0%*
 - *Conjunto de datos Economía = rico → Soporte 75%*
 - *Growth ratio = Infinito*

- *Edad > 35, Sexo = Mujer*

- *Conjunto de datos Economía = normal → Soporte 75%*
 - *Conjunto de datos Economía = rico → Soporte 50%*
 - *Growth ratio = $75/50 = 1.5$*

Patrones emergentes (*Emerging patterns*)

- Tipos de patrones emergentes
 - *Jumping emerging patterns*
 - El Growth Ratio tiene un valor de infinito
 - Muestra una enorme diferenciación entre grupos
 - *Minimal emerging patterns*
 - Un patrón se considera emergente si el valor de Growth Ratio supera un umbral
 - Patrones cuyos subpatrones no son patrones emergentes
 - *Maximal emerging patterns*
 - Un patrón se considera emergente si el valor de Growth Ratio supera un umbral
 - Patrones cuyos superpatrones no son patrones emergentes
 - *Essential jumping emerging patterns*
 - Patrones jumping cuyos subpatrones no son jumping

Patrones emergentes (Emerging patterns)

- Ejemplos de *jumping emerging pattern*
 - Pais = Suiza, Estado civil = Soltero
 - Conjunto de datos Economía = normal \rightarrow Soporte 0%
 - Conjunto de datos Economía = rico \rightarrow Soporte 75%
 - Growth ratio = Infinito ***jumping emerging pattern!***
 - Este ejemplo no es *essential jumping emerging pattern*. Si analizamos sus subpatrones:
 - Pais = Suiza
 - Growth ratio = Infinito ***jumping emerging pattern!***
 - Estado civil = Soltero
 - Growth ratio = $0.75/0.25 = 3$ ***NO es jumping emerging pattern!***

Patrones emergentes (Emerging patterns)

- Propuesta simple de extracción de patrones emergentes
 1. Dividir el conjunto de datos inicial en dos subconjuntos A y B de acuerdo con alguna propiedad de interés
 2. Aplicar un algoritmo de extracción de patrones sobre el subconjunto A
 3. Evaluar cada patrón resultante en el subconjunto B
 4. Quedarse con los patrones cuyo Growth Ratio sea mayor que un umbral
 5. Aplicar un algoritmo de extracción de patrones sobre el subconjunto B
 6. Evaluar cada patrón resultante en el subconjunto A
 7. Quedarse con los patrones cuyo Growth Ratio sea mayor que un umbral

Patrones emergentes (Emerging patterns)

- Propuesta específicas de extracción de patrones emergentes
 - Tree-based JEP
 - BCEP
 - iEPMiner
 - StrongJEP
 - Top-k minimal JEP
 - DGCP-Tree
 - CEPMine
 - EvAEP
 - etc

A stylized sunburst graphic in shades of purple and blue, located in the top-left corner of the slide. It features a semi-circle on the left with several rays extending outwards to the right.

¡Gracias!

UCO
ONLINE

A decorative horizontal bar at the bottom of the slide, consisting of alternating yellow and red rectangular segments.