

# Un Algoritmo para el Descubrimiento Automático de Prerrequisitos Curriculares.

An Algorithm for Automatic Discovery of Curricular Prerequisites.

Sara Ruiz Ruiz

Tutor: Beatriz Barros Blanco

20 de julio de 2020

## 1. Introducción

- Planteamiento del Problema
- Fases del Trabajo

## 2. Contexto del Trabajo

## 3. Algoritmo para la detección de Prerrequisitos

## 4. Aplicación del Algoritmo a datos reales

## 5. Conclusiones y propuestas de investigación futuras

# PLANTEAMIENTO DEL PROBLEMA

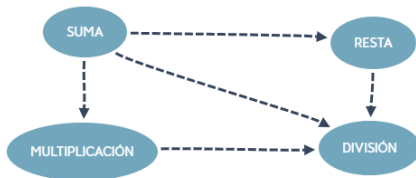
## Definición (Prerrequisito)

*Un concepto  $A$  se denomina **prerrequisito** de otro concepto  $B$  si es altamente recomendable aprender  $A$  antes que  $B$  para una buena y completa comprensión del segundo.*

# PLANTEAMIENTO DEL PROBLEMA

## Definición (Prerrequisito)

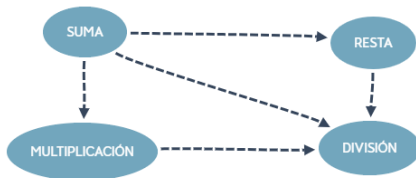
*Un concepto A se denomina **prerrequisito** de otro concepto B si es altamente recomendable aprender A antes que B para una buena y completa comprensión del segundo.*



# PLANTEAMIENTO DEL PROBLEMA

## Definición (Prerrequisito)

Un concepto *A* se denomina **prerrequisito** de otro concepto *B* si es altamente recomendable aprender *A* antes que *B* para una buena y completa comprensión del segundo.



Objetivo principal: Descubrimiento automático de prerrequisitos en un dominio de aprendizaje.

# FASES DEL TRABAJO

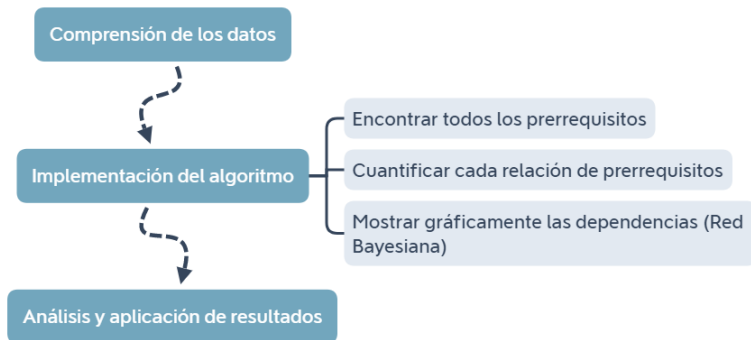


Figura: Fases del Trabajo.

# CONTEXTO DEL TRABAJO

## Representación de resultados mediante Redes Bayesianas

### Definición (Red Bayesiana)

*Una red bayesiana es una representación en forma de grafo acíclico dirigido de un modelo probabilístico.*

# CONTEXTO DEL TRABAJO

## Representación de resultados mediante Redes Bayesianas

### Definición (Red Bayesiana)

*Una red bayesiana es una representación en forma de grafo acíclico dirigido de un modelo probabilístico.*

Se compone de:

- Nodos  $\leftrightarrow$  Temas



# CONTEXTO DEL TRABAJO

## Representación de resultados mediante Redes Bayesianas

### Definición (Red Bayesiana)

*Una red bayesiana es una representación en forma de grafo acíclico dirigido de un modelo probabilístico.*

Se compone de:

- Nodos  $\leftrightarrow$  Temas
- Aristas direccionales  $\leftrightarrow$   
Relación: *Ser  
prerrequisito de*

# CONTEXTO DEL TRABAJO

## Representación de resultados mediante Redes Bayesianas

### Definición (Red Bayesiana)

*Una red bayesiana es una representación en forma de grafo acíclico dirigido de un modelo probabilístico.*

Se compone de:

- Nodos  $\leftrightarrow$  Temas
- Aristas direccionales  $\leftrightarrow$   
Relación: *Ser  
prerrequisito de*
- Tablas de probabilidades  
condicionadas (una por  
cada nodo)

## CONTEXTO DEL TRABAJO

## Representación de resultados mediante Redes Bayesianas

## Definición (Red Bayesiana)

*Una red bayesiana es una representación en forma de grafo acíclico dirigido de un modelo probabilístico.*

Se compone de:

- Nodos  $\leftrightarrow$  Temas
- Aristas direccionales  $\leftrightarrow$  Relación: *Ser prerequisite de*
- Tablas de probabilidades condicionadas (una por cada nodo)

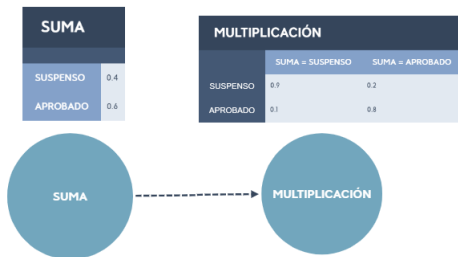
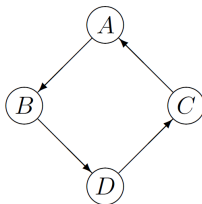


Figura: Ejemplo de una Red Bayesiana.

# CONTEXTO DEL TRABAJO

## Representación de resultados mediante Redes Bayesianas



**Figura:** Ejemplo de ciclo. NO representa una Red Bayesiana.

# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Elección de un Método Estadístico

Consideramos:

Elección (Detección de Prerrequisitos usando Probabilidades)

$T_i$  es prerrequisito de  $T_j \iff$

$$P(T_j = \text{Suspendo} \mid T_i = \text{Suspendo}) > \theta$$

donde  $\theta \in [0, 1]$  es el **Umbral del Fortaleza**.

# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Elección de un Método Estadístico

Consideramos:

Elección (Detección de Prerrequisitos usando Probabilidades)

$T_i$  es prerrequisito del  $T_j \iff$

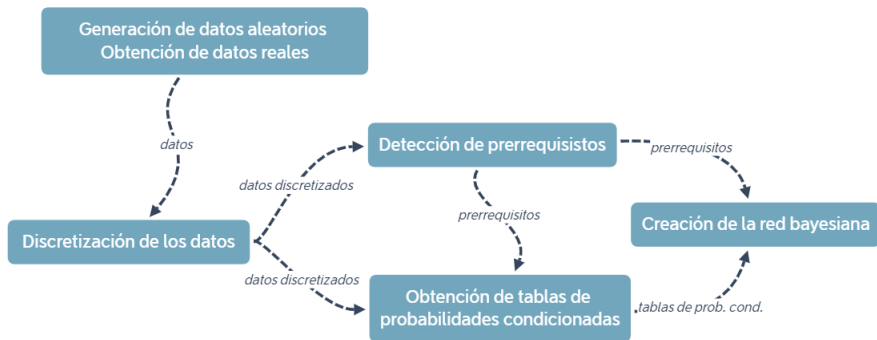
$$P(T_j = \text{Suspendo} \mid T_i = \text{Suspendo}) > \theta$$

donde  $\theta \in [0, 1]$  es el **Umbral del Fortaleza**.

Observación: Tal probabilidad **cuantifica** la relación “*Ser prerrequisito de*”

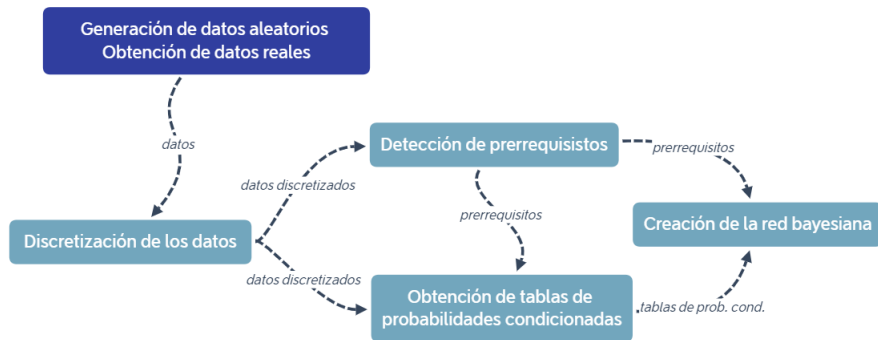
# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Etapas del Algoritmo



# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

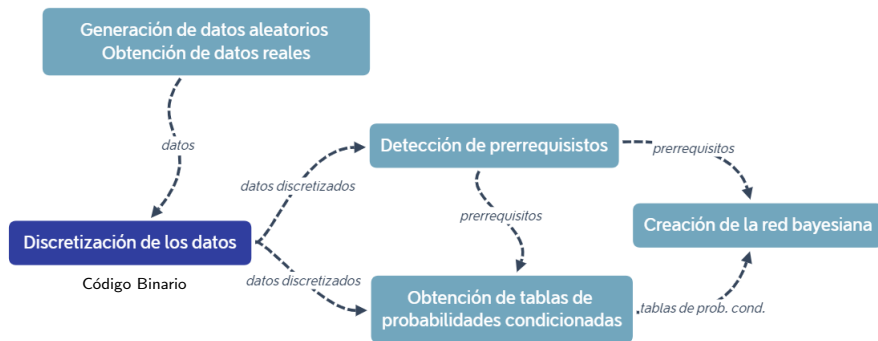
## Etapas del Algoritmo





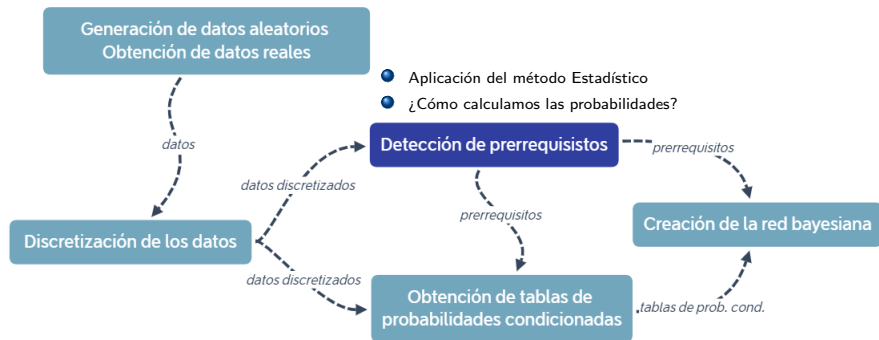
### ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Etapas del Algoritmo



# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Etapas del Algoritmo



# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

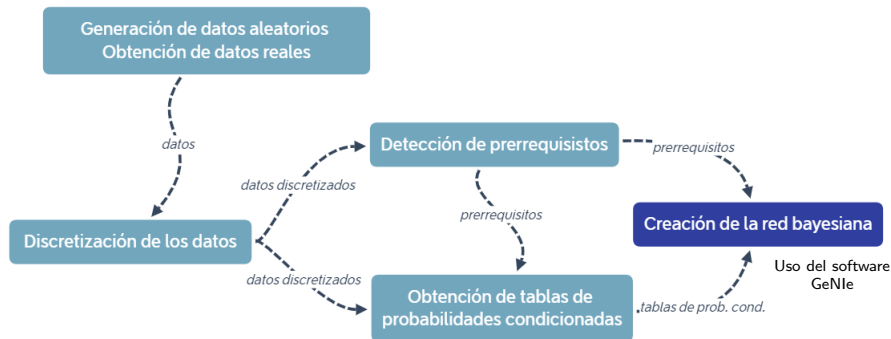
## Etapas del Algoritmo



¿Qué probabilidades son necesarias calcular?

# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Etapas del Algoritmo



# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

Errores detectados y soluciones implementadas

## 1. Posible existencia de ciclos

# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Errores detectados y soluciones implementadas

### 1. Posible existencia de ciclos

- Eliminar ciclos de dos nodos: Juntarlos en uno único.

# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Errores detectados y soluciones implementadas

### 1. Posible existencia de ciclos

- Eliminar ciclos de dos nodos: Juntarlos en uno único.
- Eliminar ciclos de más de dos nodos: Se reduce al caso anterior, gracias al teorema.

### Teorema

*Sea un grafo dirigido asociado a una distribución de probabilidad  $P$  que sigue nuestra elección estadística (esto es;  $T_i$  es padre de  $T_j \iff P(T_j = \text{Suspense} \mid T_i = \text{Suspense}) > \theta$ ). Supongamos que existe un ciclo en el grafo; entonces, el grafo presenta un ciclo de únicamente dos nodos.*

# ALGORITMO: DETECCIÓN DE PRERREQUISITOS

## Errores detectados y soluciones implementadas

### 1. Posible existencia de ciclos

- Eliminar ciclos de dos nodos: Juntarlos en uno único.
- Eliminar ciclos de más de dos nodos: Se reduce al caso anterior, gracias al teorema.

### Teorema

*Sea un grafo dirigido asociado a una distribución de probabilidad  $P$  que sigue nuestra elección estadística (esto es;  $T_i$  es padre de  $T_j \iff P(T_j = \text{Suspense} \mid T_i = \text{Suspense}) > \theta$ ). Supongamos que existe un ciclo en el grafo; entonces, el grafo presenta un ciclo de únicamente dos nodos.*

- ### 2. Posible división por cero al calcular probabilidades: Fijar los valores de las probabilidades a 0.5.



# APLICACIÓN DEL ALGORITMO A DATOS REALES

## El sistema SIETTE

**SIETTE** (Sistema Inteligente de Evaluación mediante Test para TeleEducación) es una plataforma que realiza pruebas de aprendizaje online.



# APLICACIÓN DEL ALGORITMO A DATOS REALES

## El sistema SIETTE

**SIETTE** (Sistema Inteligente de Evaluación mediante Test para TeleEducación) es una plataforma que realiza pruebas de aprendizaje online.

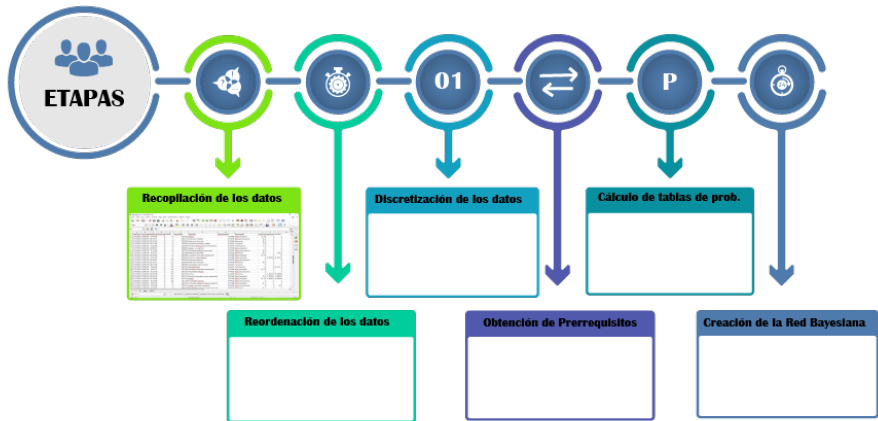


Usamos la plataforma para obtener los datos masivos para la aplicación del algoritmo a datos reales.

## APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo

Datos usados: Resultados de pruebas de alumnos de la asignatura **Estructuras Discretas**.



# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo

datosreales.xls - LibreOffice Calc

Archivo Editar Ver Insertar Formato Hoja Datos Herramientas Ventana Ayuda

Calibri 11

	A	B	C	D	E	F	G	H	I	J	K	L	M
	IDSESION	IDALUMNO	IDPREGUNTA	IDTIPOPREGUNTA	ORDEN	IDSUBTEMA	SUBTEMA	IDTEMAPADRE	TEMAPADRE	NIVELSESION	MAXPUNTO	PUNTOS	
1	376552	100019	96576	2	0	22302	grboles	18768	Teor'a de grafos	-0,15	1	1	
2	376552	100019	103471	2	1	22272	Conjuntos notables	24425	Teor'a de conjuntos	-2,4	1	1	
3	376552	100019	103733	2	2	22280	Relaciones de orden	24426	Relaciones	-0,6	1	1	
4	376552	100019	103754	2	3	22282	Funciones parciales y totales	24427	Funciones	-1,8	1	1	
5	376552	100019	96220	2	4	22289	Principios básicos de la combinatoria	20947	Combinatoria	-1,5	1	1	
6	376552	100019	103831	2	5	22686	Conexión en grafos	18768	Teor'a de grafos	-3	1	-1	
7	376552	100019	83113	4	6	18785	Conceptos básicos de conjuntos	24425	Teor'a de conjuntos	-0,9	1	1	
8	376552	100019	95357	2	7	22280	Relaciones de orden	24426	Relaciones	-0,6	1	-0,25	
9	376552	100019	103916	5	8	22298	Conceptos avanzados sobre grafos	18768	Teor'a de grafos	-0,75	1	0	
10	376552	100019	90970	4	9	22299	Caminos y conectividad	18768	Teor'a de grafos	-3	0,9999	0,3333	
11	376552	100019	103603	5	10	22269	Conjunto potencia	24425	Teor'a de conjuntos	-3	1	0	
12	376552	100019	103571	2	11	22280	Relaciones de orden	24426	Relaciones	-0,6	1	1	
13	376552	100019	103756	2	12	22284	Propiedades de funciones	24427	Funciones	-3	1	-1	
14	376552	100019	96242	2	13	22293	Permutaciones	20947	Combinatoria	-3	1	-0,1111	
15	376552	100019	96571	2	14	22298	Conceptos avanzados sobre grafos	18768	Teor'a de grafos	-0,75	1	1	
16	376552	100019	83332	2	15	19126	Propiedades básicas	24425	Teor'a de conjuntos	3	1	1	
17	376552	100019	103918	4	16	22278	Cierres	24426	Relaciones	-3	1,0002	0,6668	
18	376552	100019	103912	4	17	22301	Recorrido de grafos y tipos especiales	18768	Teor'a de grafos	2,4	1,0003	1,0003	
19	376552	100019	103590	4	18	22302	grboles	18768	Teor'a de grafos	-0,15	1,0002	0,6668	
20	376552	100019	83339	2	19	19126	Propiedades básicas	24425	Teor'a de conjuntos	3	1	1	
21	376554	100019	96570	2	0	22295	Conceptos básicos de teor'a de grafos	18768	Teor'a de grafos	-2,7	1	1	
22	376554	100019	95332	2	1	22270	Tuplas y producto cartesiano	24425	Teor'a de conjuntos	-3	1	-0,2	
23	376554	100019	95353	2	2	24426	Relaciones	24426	Relaciones	-3	1	1	

Buscar Buscar todo Cadena formateada Distinguir mayúsculas y minúsculas

Hoja 1 de 2 PageStyle: Datos Promedio: Suma: 0 100 %

**ETAPAS**

**Recopilación de los datos**

**Reordenación de los datos**

**Discretización de los datos**

**Obtención de Prerrequisitos**

**Cálculo de tablas de prob.**

**Creación de la Red Bayesiana**

# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo

ordenadoreal.xlsx - LibreOffice Calc

Archivo Editar Ver Insertar Formato Hoja Datos Herramientas Ventana Ayuda

Calibri 11

AC44

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1		18768	24425		24427	20947															
2	100019	0,565505	0,565557	0,432879	0,222225	0,424542															
3	101247	0,614384	0,792265	0,739584	0,78125	0,459663															
4	100109	0,961133	1	0,916679	0,773807	0,922621															
5	100386	0,834351	0,65834	0,75625	0,749996	0,631792															
6	100342	0,055552	0,333333	0,534725	0,666667	0,238417															
7	101218	0,753516	0,800142	0,829423	0,709565	0,838974															
8	100598	1,000029	1	1	1,0001	1															
9	101012	0,880967	0,866667	1	0,666667	1															
10	100818	-0,00256	0,5	0,916675	0,45835	0,13255															
11	99904	0,669664	0,98334	0,593763	0,479175	0,875															
12	99815	0,51749	0,728577	0,605564	0,785714	0,531921															
13	100604	0,88572	1,000008	0,95001	1	0,8															
14	100790	0,971633	0,825645	0,9859	0,807692	0,831046															
15	99887	1,000043	1,00004	1	1	1															
16	101244	0,420643	0,833333	0,5625	-0,45833	0,81945															
17	100572	-0,14286	1	1	0	-0,1111															
18	100672	0,88099	0,977787	0,911117	1	0,8125															
19	100980	0,587302	0,6889	0,689592	0,166667	0,909725															
20	101102	1,000014	1,00004	1	1	1															
21	100918	0,418048	0,786227	0,486398	0,169447	0,70681															
22	101066	0,956375	0,877787	0,794461	0,740744	0,75															
23	100954	0,839437	0,720222	0,836619	0,428867	0,6277															
24	100932	0,071443	0,73334	1	-0,125	-0,15655															
25	100912	0,693479	0,810719	0,748184	0,859375	0,579256															

Sheet1 Sheet2

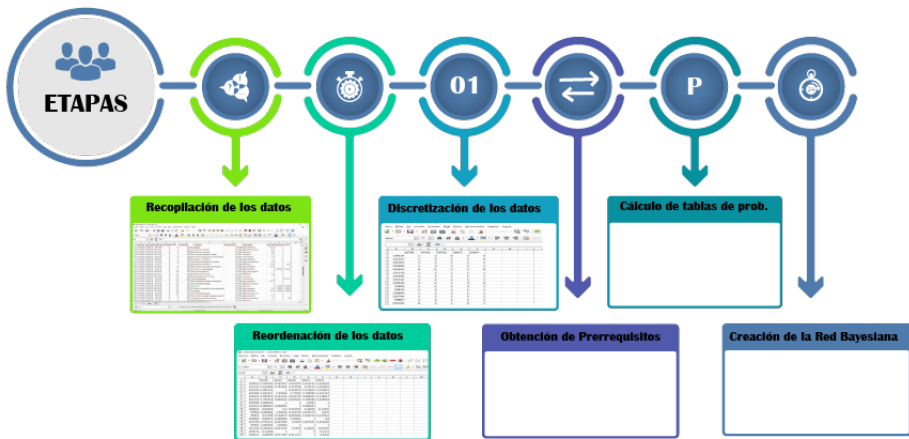
Buscar

Buscar todo ☐ Cadena formateada ☐ Distinguir mayúsculas y minúsculas

Hoja 1 de 2 PageStyle\_Sheet1 Promedio: Suma: 0 100 %

# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo



# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo

discretizadoreal.xlsx - LibreOffice Calc

Archivo Editar Ver Insertar Formato Hoja Datos Herramientas Ventana Ayuda

Calibri 11

R25

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1		18768	24425	24426	24427	20947															
2	100019	1	1	0	0	0															
3	101247	1	1	1	1	0															
4	100109	1	1	1	1	1															
5	100386	1	1	1	1	1															
6	100342	0	0	1	1	0															
7	101218	1	1	1	1	1															
8	100598	1	1	1	1	1															
9	101012	1	1	1	1	1															
10	100818	0	0	1	0	0															
11	99904	1	1	1	0	1															
12	99815	1	1	1	1	1															
13	100604	1	1	1	1	1															
14	100790	1	1	1	1	1															
15	99887	1	1	1	1	1															
16	101244	0	1	1	0	1															
17	100572	0	1	1	0	0															
18	100672	1	1	1	1	1															
19	100980	1	1	1	0	1															
20	101102	1	1	1	1	1															
21	100918	0	1	0	0	1															
22	101066	1	1	1	1	1															
23	100954	1	1	1	0	1															
24	100932	0	1	1	0	0															
25	100912	1	1	1	1	1															

Sheet1

Buscar

Buscar todo ☐ Cadena formateada ☐ Distinguir mayúsculas y minúsculas

Hoja 1 de 1

PageStyle\_Sheet1

Promedio: Suma: 0

100 %



**ETAPAS**

**Recopilación de los datos**

**Discretización de los datos**

**Cálculo de tablas de prob.**

**Obtención de Prerrequisitos**

**Reordenación de los datos**

**Creación de la Red Bayesiana**

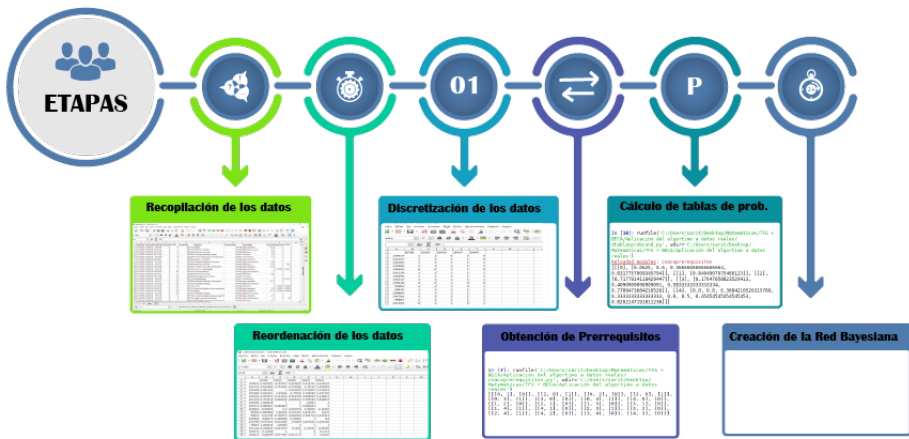
# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo

```
In [7]: runfile('C:/Users/sarit/Desktop/Matemáticas/TFG +  
BECA/Aplicación del algoritmo a datos reales/  
csacoprerequisitos.py', wdir='C:/Users/sarit/Desktop/  
Matemáticas/TFG + BECA/Aplicación del algoritmo a datos  
reales')  
[[[0, 1], [0]], [[1, 0], [1]], [[0, 2], [0]], [[2, 0], [1]],  
[[0, 3], [1]], [[3, 0], [0]], [[0, 4], [1]], [[4, 0], [0]],  
[[1, 2], [0]], [[2, 1], [0]], [[1, 3], [0]], [[3, 1], [0]],  
[[1, 4], [1]], [[4, 1], [0]], [[2, 3], [1]], [[3, 2], [0]],  
[[2, 4], [1]], [[4, 2], [0]], [[3, 4], [0]], [[4, 3], [0]]]
```

# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo



# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo

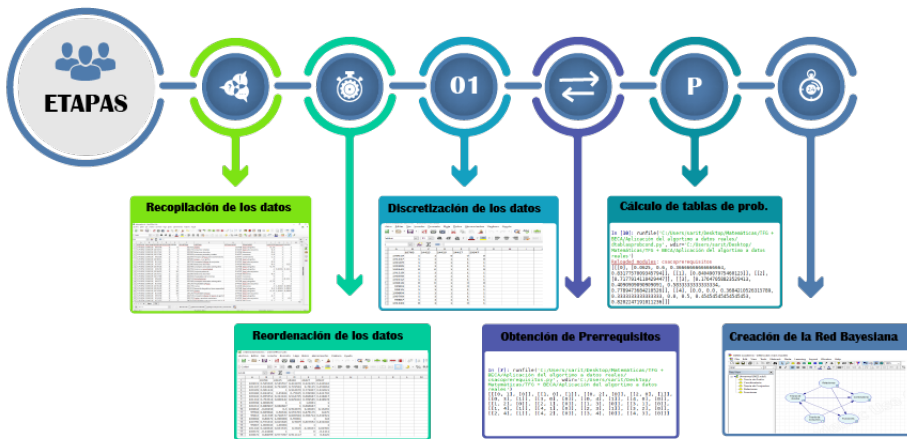
```
In [10]: runfile('C:/Users/sarit/Desktop/Matemáticas/TFG +
BECA/Aplicación del algoritmo a datos reales/
dtablasprobcond.py', wdir='C:/Users/sarit/Desktop/
Matemáticas/TFG + BECA/Aplicación del algoritmo a datos
reales')
```

Reloaded modules: csacoprerequisitos

```
[[[0], [0.0625, 0.6, 0.36666666666666664,
0.8317757009345794]], [[1], [0.8404907975460123]], [[2],
[0.7177914110429447]], [[3], [0.17647058823529413,
0.409090909090909091, 0.5833333333333334,
0.7789473684210526]], [[4], [0.0, 0.0, 0.3684210526315789,
0.3333333333333333, 0.0, 0.5, 0.45454545454545453,
0.8202247191011236]]]
```

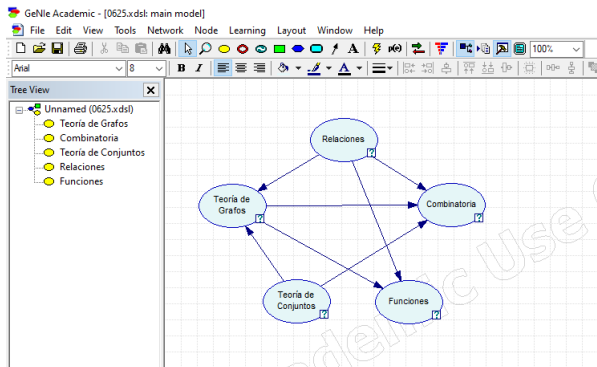
# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo



## APLICACIÓN DEL ALGORITMO A DATOS REALES

## Etapas del Algoritmo



# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Relaciones entre los temas padre

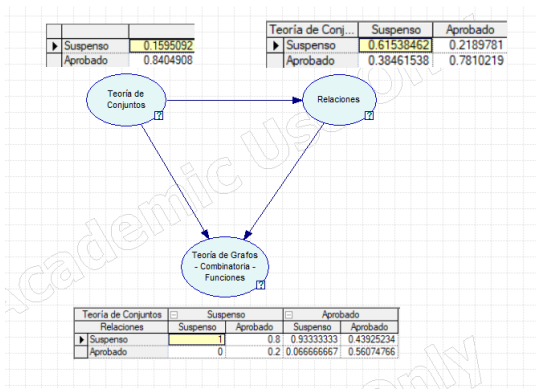
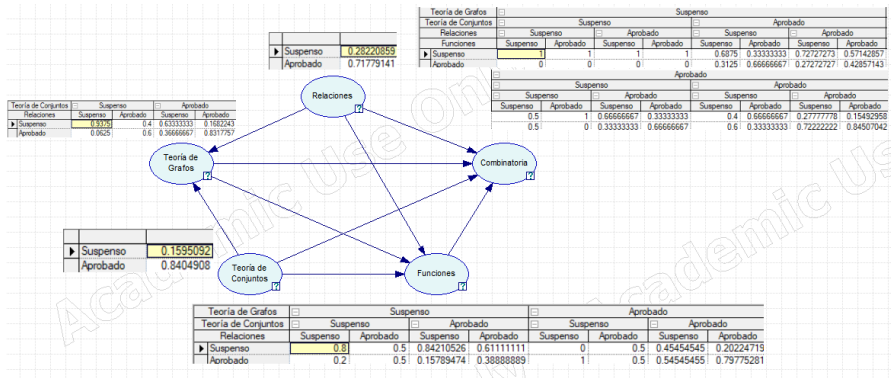


Figura: Red Bayesiana obtenida con  $\theta = 0,60$ .

Figura: Red Bayesiana obtenida con  $\theta = 0,625$ .



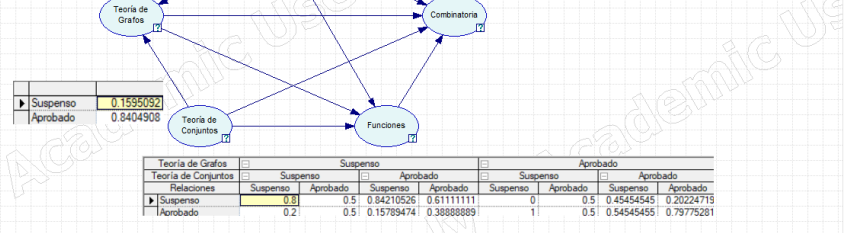


Figura: Red Bayesiana obtenida con  $\theta = 0,625$ .

Secuencia óptima: **Teoría de Conjuntos - Relaciones - Teoría de Grafos - Funciones - Combinatoria**

# APLICACIÓN DEL ALGORITMO A DATOS REALES

Comparación con un modelo de experto

Secuencia del algoritmo:

**Teoría de Conjuntos**

**Relaciones**

**Teoría de Grafos**

**Funciones**

**Combinatoria**

# APLICACIÓN DEL ALGORITMO A DATOS REALES

Comparación con un modelo de experto

Secuencia del algoritmo:

**Teoría de Conjuntos**  
**Relaciones**  
**Teoría de Grafos**  
**Funciones**  
**Combinatoria**

Secuencia en SIETTE:

**Teoría de Conjuntos**  
**Relaciones**  
**Funciones**  
**Combinatoria**  
**Teoría de Grafos**

# APLICACIÓN DEL ALGORITMO A DATOS REALES

## Comparación con un modelo de experto

Secuencia del algoritmo:

**Teoría de Conjuntos**

**Relaciones**

**Teoría de Grafos**

**Funciones**

**Combinatoria**

Secuencia en SIETTE:

**Teoría de Conjuntos**

**Relaciones**

**Funciones**

**Combinatoria**

**Teoría de Grafos**

# CONCLUSIONES

1. Se ha profundizado en el concepto de *Prerrequisito*.

# CONCLUSIONES

1. Se ha profundizado en el concepto de *Prerrequisito*.
2. Se ha desarrollado un algoritmo desde cero.
  - Apto para cualquier conjunto de datos.
  - Con las dependencias encontradas cuantificadas.
  - Con fácil interpretación de los resultados.

# CONCLUSIONES

1. Se ha profundizado en el concepto de *Prerrequisito*.
2. Se ha desarrollado un algoritmo desde cero.
  - Apto para cualquier conjunto de datos.
  - Con las dependencias encontradas cuantificadas.
  - Con fácil interpretación de los resultados.
3. Se ha implementado el algoritmo en *Python*.

# CONCLUSIONES

1. Se ha profundizado en el concepto de *Prerrequisito*.
2. Se ha desarrollado un algoritmo desde cero.
  - Apto para cualquier conjunto de datos.
  - Con las dependencias encontradas cuantificadas.
  - Con fácil interpretación de los resultados.
3. Se ha implementado el algoritmo en *Python*.
4. Se ha aplicado a datos reales y analizado sus resultados.



# POSIBLES VÍAS DE FUTURO AVANCE

1. Posibilidad de trabajar con rangos de valores.

# POSIBLES VÍAS DE FUTURO AVANCE

1. Posibilidad de trabajar con rangos de valores.
2. Integración del algoritmo en el sistema SIETTE.

# POSIBLES VÍAS DE FUTURO AVANCE

1. Posibilidad de trabajar con rangos de valores.
2. Integración del algoritmo en el sistema SIETTE.
3. Comparación del algoritmo con otros métodos.

GRACIAS POR SU ATENCIÓN

GRACIAS POR SU ATENCIÓN

