Detección de tópicos en noticias del diario **La Nación**

(Con herramientas de análisis de redes complejas, obvio...)



Grupo:

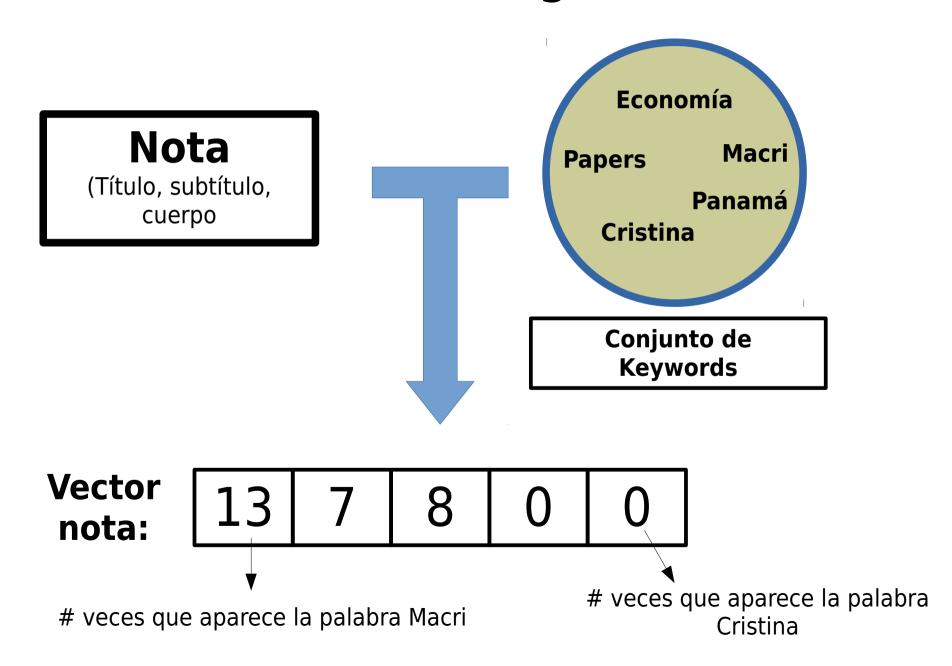
Raúl Barriga. Mariela Celis. Jimmy Masías. Sebastián Pinto.

Datos y objetivos

 <u>Datos</u>: notas de la sección política del diario La Nación (edición impresa): 05-08 al 13-10-2016.

 Objetivos: detectar los tópicos que cubrió el diario durante esa ventana temporal.

Metodología



Metodología

Keywords

 Notas
 10
 10

 0
 1
 1
 5
 10

 0
 15
 16
 0
 0

 11
 0
 0
 12
 2

Matriz notas x keywords

Metodología

- Reducción de la dimensión (PCA): quita redundancia.
 Ej: Panama, Papers → única keywords. Notas parecidas tienden a ser mas parecidas, pero hay perdida de información.
- Definición de similaridad (correlación) entre notas → matriz de similaridad.

Matriz notas x notas

1	0.9	0.1
0.9	1	0.2
0.1	0.2	1

Similaridad entre

nota 1 y 2

Punto de partida y objetivos

 Tomar la matriz de similaridad como una matriz de adyacencia de un grafo completo, con links pesados.

- Detección de comunas: tópicos principales.
- <u>Network priorization</u>: detección de tópicos a partir de información parcial.
- Evolución temporal del grafo.