

## **Tarea Unidad 4 INFO343**

### **Aplicación de Modelos de Tópicos Dinámicos**

Cada estudiante escoge uno de los archivos con enlaces a noticias de un país latinoamericano, referidas a la epidemia, y que extienden entre enero y junio del 2020.

**Objetivo:** Analizar la evolución de las noticias de la pandemia en un país de Latinoamérica, extrayendo del conjunto de datos los distintos tópicos presentes en las mismas y la evolución en la aparición en las noticias de los principales conceptos de cada tópico, a lo largo de los meses estudiados.

#### **Actividades:**

- 1) Procesar el archivo con los enlaces a noticias para rescatar los textos de cada una (cuando sea posible)
- 2) Realizar el pre-procesamiento de las noticias hasta su representación por palabras clave. Definir las slices para el análisis de tópicos dinámicos. Normalmente una por mes.
- 3) Realizar el análisis de tópicos dinámicos, incluyendo calidad de los tópicos y evolución en el tiempo de los conceptos principales.
- 4) Redactar un informe con los principales elementos de la metodología utilizada, resultados y conclusiones.

#### **Fechas de Entregas:**

##### **Primer Avance: miércoles 5 de agosto**

Informe de las actividades (1) y (2) incluyendo una estadística preliminar de los datos: Número total de noticias, tamaño del vocabulario de palabras clave obtenido, número de noticias por slice, anexo con el archivo de noticias depurado.

**Informe final: lunes 17 de agosto.**