

Introducción a la Teoría de Grafos

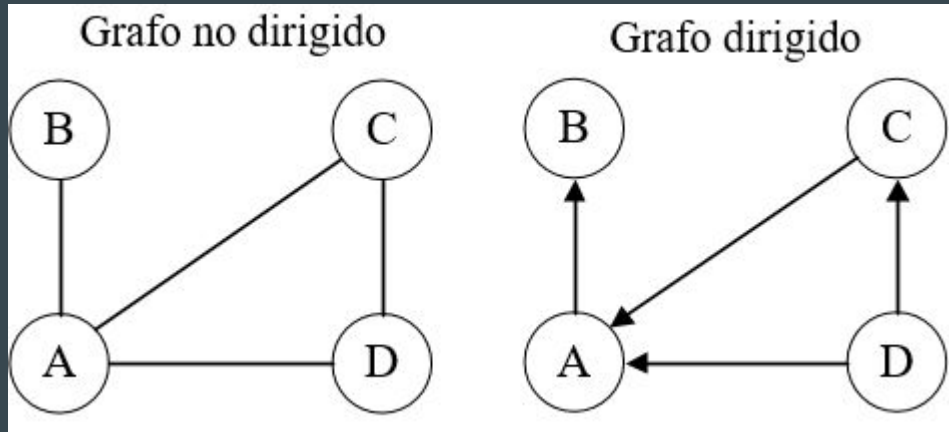
...

KeepCoding - Bootcamp de Big Data & Machine Learning

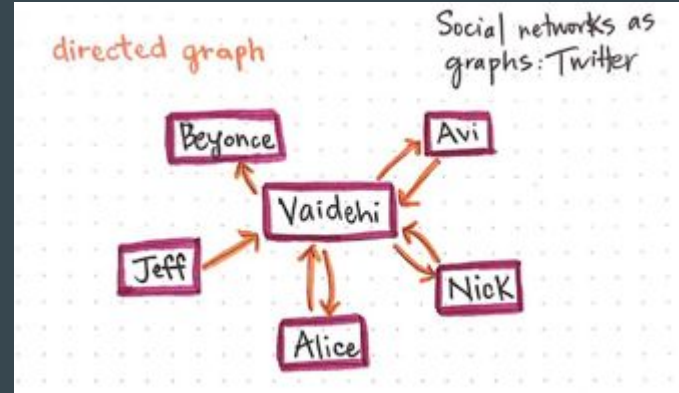
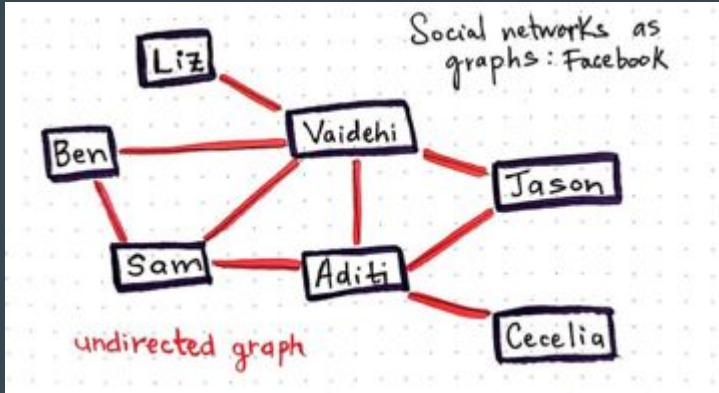
Índice

1. Tipos de grafos
2. Centralidad
3. Modularidad
4. Ejemplos

1. Tipos de grafos



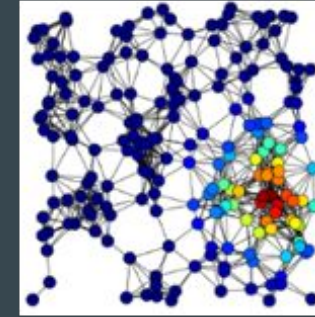
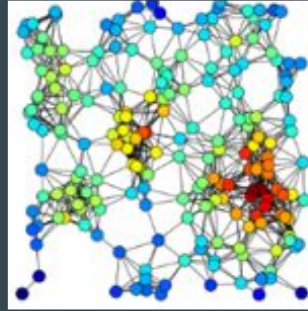
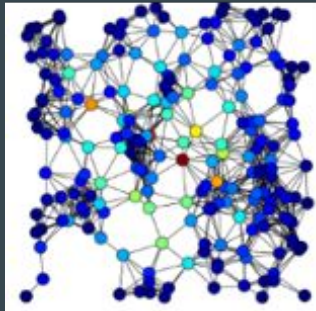
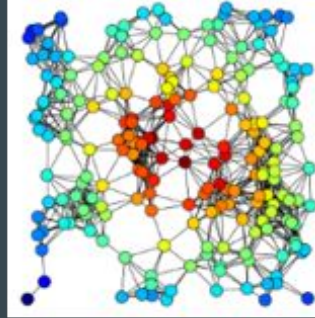
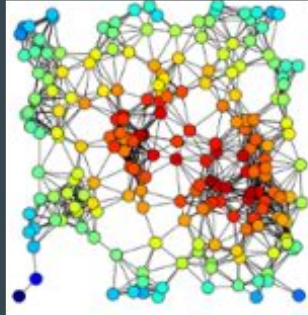
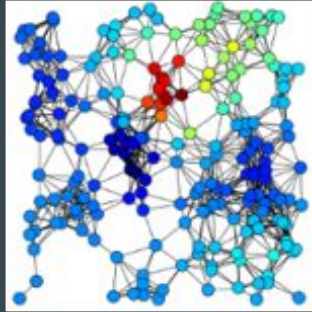
1. Tipos de grafos



2. Centralidad

- Objetivo: encontrar los **nodos más relevantes** en una red
- La relevancia - o influencia - de un nodo en una red viene determinada por cómo se **encuentra conectado** ese nodo en esa red
- Ejemplo: influencers en Twitter (número de seguidores, número de retweets, número de favs, número de seguidores de sus seguidores, etc.)

2. Centralidad



2. Centralidad - Métricas

- **Degree** (grado)
 - Número de enlaces que tiene un nodo con los demás
- **Closeness** (cercanía)
 - Se basa en la cercanía de un nodo con respecto al resto de nodos, comparado con los demás
- **Betweenness** (intermediación)
 - Número de veces que un nodo se encuentra en el camino de mínima distancia entre otros nodos
- **Eigencentrality** (autovector)
 - Similar al degree, pero teniendo en cuenta también el degree de sus vecinos

2. Centralidad - Métricas

- Degree (grado)
 - Número de enlaces que tiene un nodo con los demás
 - Útil para encontrar individuos altamente conectados
 - Medida más básica de centralidad pero útil en general
 - Si el grafo es dirigido:
 - in-degree: n° enlaces entrantes
 - out-degree: n° enlaces salientes

2. Centralidad - Métricas

- **Closeness** (cercanía)
 - Se basa en la **cercanía** de un nodo con respecto al resto de nodos, **comparado con los demás**
 - Útil para encontrar **individuos bien situados** que, potencialmente, podrían **influir** en la red de manera más rápida
 - En redes grandes con muchas conexiones pierde un poco de sentido
 - Muy útil si se mide no en la red completa si no en comunidades

2. Centralidad - Métricas

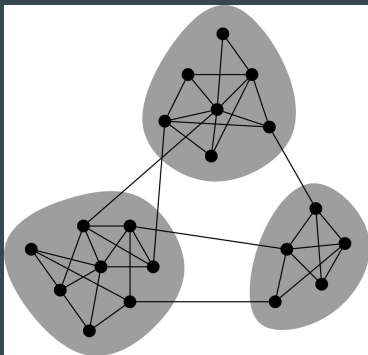
- **Betweenness** (intermediación)
 - Número de veces que un nodo se encuentra en el camino de mínima distancia entre otros nodos
 - Permite encontrar individuos con - potencialmente - **poder de influencia** a lo largo de toda la red
 - Valores altos pueden indicar que nos encontramos con “**nodos puente**” o “**nodos periféricos**”

2. Centralidad - Métricas

- **Eigencentrality** (autovector)
 - Similar al degree pero **teniendo en cuenta el degree de sus vecinos**
 - Útil detectando *influencers*
 - Útil para detectar individuos potencialmente influyentes en el total de la red
 - Ejemplo: PageRank
 - Algoritmo creado por Google en 1999 usado en el ranking de páginas web
 - Tiene en cuenta también la dirección de los enlaces y su peso

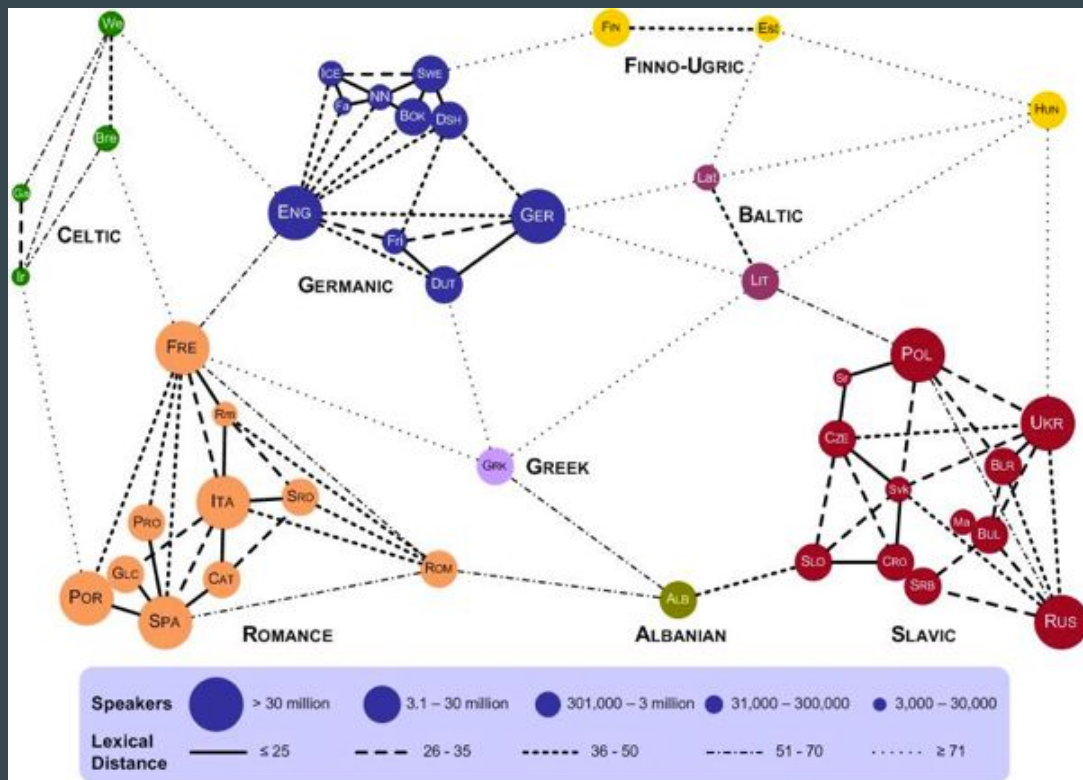
3. Modularidad

- Mide la **densidad** en una red o grafo
- Permite detectar nodos que se encuentran más densamente conectados entre sí, comparados con el resto de la red
- Muy útil para **encontrar comunidades**



4. Ejemplos

- Similitud entre diferentes idiomas



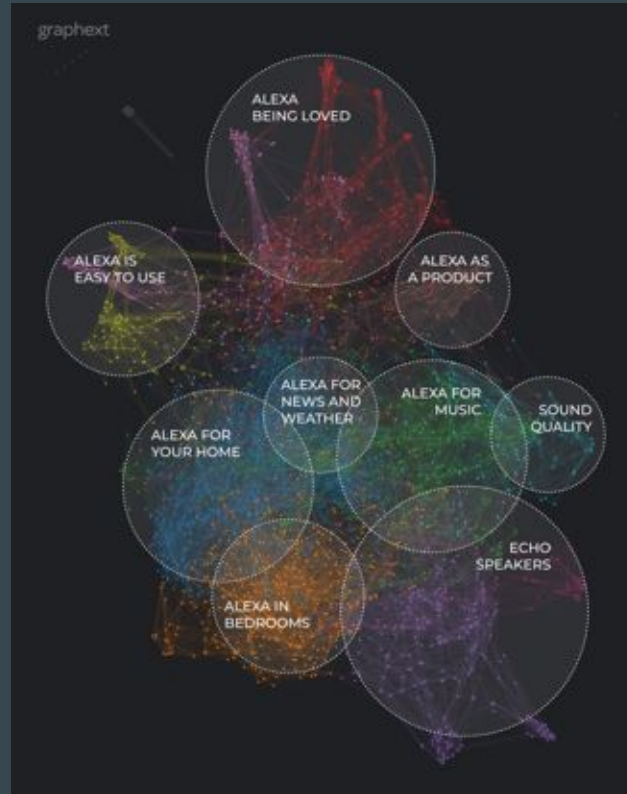
4. Ejemplos

- Coocurrencias de personajes en Game of Thrones



4. Ejemplos

- Peticiones de usuarios a Alexa



¡Vamos al lío!