



Fase 0: Cargue de datos y análisis visual de la red del proyecto en notebook o Gephi

Tabla de contenido

OBJETIVOS	2
CONTEXTO	3
DATOS	3
METODOLOGÍA	4
ENTREGABLES	6
EVALUACIÓN.....	6
REFERENCIAS	6

Objetivos

El objetivo general del proyecto es analizar las redes sociales de diferentes inversionistas para caracterizar su comportamiento e identificar los atributos significativos a la hora de financiar compañías. Se espera que los estudiantes sean capaces de aplicar los conocimientos vistos en el curso en un contexto real y llevar a cabo análisis e interpretaciones en el contexto del problema dado, utilizando las herramientas adecuadas.

La siguiente lectura tiene como propósito presentar la **Fase 0: Cargue de datos y análisis visual de la red del proyecto en notebook o Gephi**. En primer lugar, se presenta el contexto en el que se desarrolla el proyecto, luego se describe la base de datos con la que se trabajará, después se presentan la metodología y entregables y, finalmente, la evaluación de esta fase.

Objetivos del curso relacionados con esta entrega:

1. Identificar la estructura de una base de datos para el análisis de redes.
2. Reconocer y calcular medidas de centralidad y prestigio de redes sociales.

3. Interpretar los resultados de las medidas de centralidad y prestigio de redes sociales.
4. Reconocer las características de los algoritmos de visualización de redes.
5. Interpretar características de las redes a partir de su visualización.

Contexto

Normalmente las compañías buscan fuentes de financiación y dedican esfuerzos a mejorar sus posibilidades de recibir inversiones (Liang & Yuan, 2016). Uno de los problemas que surgen en esta búsqueda de fuentes de financiación es que existen asimetrías de información. Esto ocurre porque las compañías que buscan financiación tienen información acerca de su potencial y de sus oportunidades que los inversionistas no tienen (Shane & Cable, 2002). En este contexto el análisis de redes sociales es una herramienta que permite analizar el comportamiento de los inversionistas teniendo en cuenta sus relaciones con otros. Esto permite conocer lo que los inversionistas están buscando y los factores que determinan su comportamiento al momento de financiar empresas (Liang & Yuan, 2016).

Datos

Para lograr el objetivo del proyecto ustedes cuentan con datos de Crunchbase que contienen la información de un conjunto de organizaciones y sus inversionistas. Los datos están organizados en los siguientes archivos de Excel de la siguiente manera:

- **Base inicial:** este archivo contiene todos los atributos de los nodos de la red que se va a analizar. Un nodo es una organización y cuenta con los siguientes atributos:
 - **Id:** Nombre de la empresa.
 - **Location:** Ubicación de la empresa.
 - **Operating Status:** Estado de operación de la empresa. Puede estar Activa (active) o Cerrada (closed).
 - **Company Type:** Indica si la organización es sin ánimo de lucro (non profit) o con ánimo de lucro (profit).
 - **Number of Investments:** Cantidad total de inversiones realizadas.
 - **Number of Lead Investments:** Cantidad de inversiones realizadas como inversor principal.
 - **Number of Diversity Investments:** Cantidad de inversiones diversificadas realizadas. Cuando se realiza una inversión diversificada se destina el dinero a invertir en diferentes portafolios.
 - **Number of Exits:** Cantidad de veces que en la compañía se han emitido acciones.

- **Number of Lead Investors:** Cantidad de inversionistas principales en la compañía.
- **Number of Investors:** Cantidad de inversionistas en la compañía.
- **Industry Groups:** Grupos de industrias a los cuales pertenecen las organizaciones. La lista de grupos de industrias que podría haber y sus industrias puede [encontrarse aquí.](#)
- **Industries:** Industrias a las cuales pertenecen las organizaciones. La lista de grupos de industrias que podría haber y sus industrias puede [encontrarse aquí.](#)
- **Number of Funding Rounds:** Cantidad de rondas de financiación de la empresa.
- **Funding Status:** Estado de financiación de la empresa. Puede ser Early Stage Venture, Late State Venture, IPO, M&A, Seed o Private Equity.
- **Last Funding Date:** Última fecha de financiación.
- **Last Funding Amount Currency (in USD):** Valor de la última financiación en dólares estadounidenses.
- **Last Funding Type:** Tipo de la última financiación. Los diferentes tipos de financiación pueden [encontrarse aquí.](#)
- **Last Equity Funding Type:** Tipo de la última financiación excluyendo la deuda. Los diferentes tipos de financiación pueden [encontrarse aquí.](#)
- **Last Equity Funding Amount Currency (in USD):** Valor de la última financiación en dólares estadounidenses excluyendo la deuda.
- **Total Funding Amount Currency (in USD):** Valor total de dinero recaudado en dólares estadounidenses durante todas las rondas de inversión.
- **Number of Events:** Cantidad total de eventos en los que la organización ha aparecido.
- **SEMrush – Monthly Visits:** Cantidad de visitas al sitio web de la compañía en el último mes incluyendo visitas al sitio móvil y desktop.
- **SEMrush - Average Visits (6 months):** Cantidad promedio de visitas mensuales al sitio web de la compañía en los últimos 6 meses incluyendo visitas al sitio móvil y desktop.
- **SEMrush – Visit Duration:** Tiempo promedio que un usuario permanece en la página web de la compañía por visita. Se mide en segundos.
- **Aberdeen - IT Spend Currency (in USD):** Monto aproximado en dólares que la compañía invierte anualmente en tecnologías de información (TI).
- **Principales inversionistas:** Nombres de los principales inversionistas de la organización.

El archivo puede descargarse en el siguiente enlace: <https://www.coursera.org/learn/analytics-redes-sociales/resources/XOI9m>

Metodología

Supongan que ustedes son los responsables en su organización de tomar los datos de Crunchbase y prepararlos para llevar a cabo un análisis visual y descriptivo utilizando el análisis de redes sociales. De

acuerdo con esto, se le encomendó la tarea de preparar, cargar los datos y realizar un análisis visual preliminar de la red utilizando Gephi y/o un notebook.

En esta fase se deberán realizar:

1. **La preparación de los datos para analizarlos con redes sociales:** Como bien saben, para analizar datos de redes sociales se requiere tener un archivo de nodos donde se encuentran los atributos de cada nodo y un archivo de arcos donde se conocen las relaciones entre cada par de nodos. Un analista de la organización les hizo entrega del archivo “Nodos” y ustedes identificaron que tienen diferentes atributos para las organizaciones, pero aún no cuentan con el archivo de arcos para llevar a cabo el análisis de redes sociales. De acuerdo con lo anterior, deben preparar el archivo de Arcos para poder seguir adelante en el proceso. Para llevar a cabo esta preparación, les brindamos las siguientes recomendaciones:

- a. En el archivo de nodos encontrarán una variable llamada “Principales inversionistas”. Esta variable contiene los nombres de los principales inversionistas en de la organización.
- b. Una relación entre un par de nodos se forma cuando una organización invierte en la otra. En ese sentido, esta red es dirigida y en el archivo de arcos deberán tener dos variables:
 - i. Source: corresponde al nombre de la organización inversionista. Es decir, la organización que invierte en la otra.
 - ii. Target: corresponde al nombre de la organización en la cual se invierte.

Esto significa que la variable “Principales inversionistas” contiene los nombres de los nodos (“Source”) que apuntan a cada nodo (“Target”) del archivo de Nodos en la red. Tengan en cuenta que solo deben crear las relaciones entre los nodos que cuentan con atributos en la base de datos.

2. **La carga de los datos en un notebook o en Gephi:** una vez tengan un archivo de nodos y un archivo de arcos podrán cargarlos en un notebook o en Gephi para proceder con el análisis.
3. **Un análisis visual de la red:** deberán identificar los actores relevantes de la red haciendo uso de las herramientas dadas en el curso hasta el momento.

Entregables

El entregable de esta fase es una infografía. Cada grupo debe:

1. Diseñar una infografía
 - a. Una o varias imágenes de la red descritas.
 - b. Un párrafo de máximo 500 palabras que contenga un análisis con los insights más importantes obtenidos hasta el momento.
2. Subir una infografía al Padlet el cual encuentran en plataforma.

Nota: Puede ver un ejemplo de la infografía dentro del padlet cuyo enlace está en la actividad.

Deben hacer uso de las herramientas vistas en el curso hasta el momento: visualización de redes y medidas de centralidad y prestigio.

Es importante tener en cuenta que pueden incluir tantas imágenes como consideren en la infografía. Cuando se utilizan imágenes de redes se hace una breve descripción de la imagen explicando elementos como: qué es un nodo, qué es un arco, qué representa el tamaño de los nodos y/o arcos y qué representa el color de los nodos y/o arcos.

Evaluación

Esta fase no tiene nota pero su presentación es obligatoria para avanzar en el proyecto.

Referencias

Liang, Y. E., & Yuan, S. T. D. (2016). Predicting investor funding behavior using crunchbase social network features. *Internet Research*, 26(1), 74–100. <https://doi.org/10.1108/INTR-09-2014-0231/FULL/XML>

Shane, S., & Cable, D. (2002). Network Ties, Reputation, and the Financing of New Ventures. *Management Science*, 48(3), 364–381. <http://www.jstor.org/stable/822571>

