



Fase 1: Cargue de datos, análisis visual de la red e identificación de actores relevantes

Tabla de contenido

OBJETIVOS	2
CONTEXTO	3
DATOS	3
METODOLOGÍA Y ENTREGABLES	5
EVALUACIÓN.....	5
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 1: C Cargue de datos, análisis visual de la red e identificación de actores relevantes**. 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 las fases y entregables y, finalmente, le evaluación de cada fase.

El proyecto se lleva a cabo en grupos cuya conformación es definida al inicio del curso por parte del equipo docente. En los contenidos del curso podrán encontrar la asignación de grupos.

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.
6. Reconocer medidas de transitividad que permiten evaluar la presencia de clusters en la red.

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.

Descargar los archivos **Nodos** y **Arcos** que encuentran en el siguiente enlace <https://www.coursera.org/learn/analytics-redes-sociales/resources/Rk5wl>

Metodología y Entregables

Estamos en la primera fase del proyecto y ustedes actuarán bajo su rol original de estudiantes de un curso de análisis de redes sociales. Esta fase que consiste en la carga de datos, análisis visual de la red e identificación de los actores relevantes.

Para la Semana 3 se espera que:

1. **Utilicen** las herramientas dadas en el curso hasta el momento (análisis visual, medidas de centralidad y medidas de transitividad).
1. **Construyan** un informe de máximo 5 páginas donde:
 - a. Se documente el proceso de carga de los datos.
 - b. Se realice un análisis visual de la red y se identifiquen los actores relevantes de la red. Para esto se espera que utilicen el trabajo realizado en la fase 0 y las medidas de **centralidad y transitividad** que consideren relevantes para realizar el análisis visual e identificar los actores relevantes. Para cada medida calculada y analizada se debe justificar su selección y relevancia en el problema.
 - c. Se debe presentar una conclusión al respecto en el contexto del problema y recomendaciones preliminares.

Evaluación

La primera fase del proyecto se evaluará a través de la siguiente matriz de calificación:

Criterio	Puntos en la nota de la fase
Se documenta el proceso de carga de los datos describiendo en detalle los pasos que se siguieron para tratar y cargar los datos antes de visualizarlos.	20

Se realiza el análisis visual de la red y se identifican sus actores relevantes. Para ello, se presentan imágenes de la red debidamente descritas y se identifican visual y analíticamente los actores relevantes de la red. En el análisis se argumenta la elección de diferentes medidas de centralidad y transitividad y cómo estas permiten identificar a los actores relevantes en la red.	50
Se realiza una conclusión sobre los resultados obtenidos y recomendaciones preliminares en el contexto del problema.	30

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>