

AMOR U ODIO ¿POR QUÉ NOS GUSTA UNA PELÍCULA?

Marcadores léxicos de éxito o fracaso a
partir de técnicas de twitter scraping

Leonardo Rubio & Angela Villate

Noviembre de 2020



OBJETIVOS

- Nos propusimos aplicar técnicas de procesamiento de lenguaje natural para encontrar patrones de las personas sobre el gusto de las películas comerciales.
- Aplicaciones que esperan decantar resultados que indiquen:
- Cómo y por qué es útil aplicar técnicas de scraping y machine learning a fenómenos usos del lenguaje común.
- Factores que explican el éxito o el fracaso de una película.



OBJETIVOS



- Perfilamiento de textos
- Análisis de datos
- Web scrapping
- Nubes de palabras
- Clasificación K-means
- Twitter scrapping
- Análisis de sentimientos





Perfilamiento de los datos



¿Dinero?

El dinero invertido en la producción (o recolectado) no influye al momento de decidir si gusta o no una película

¿Genero?

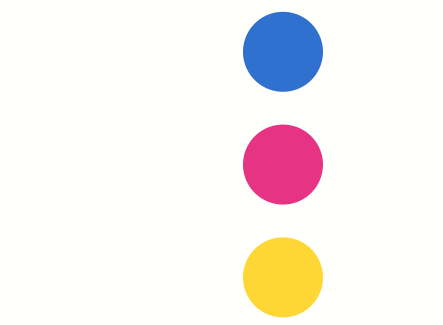
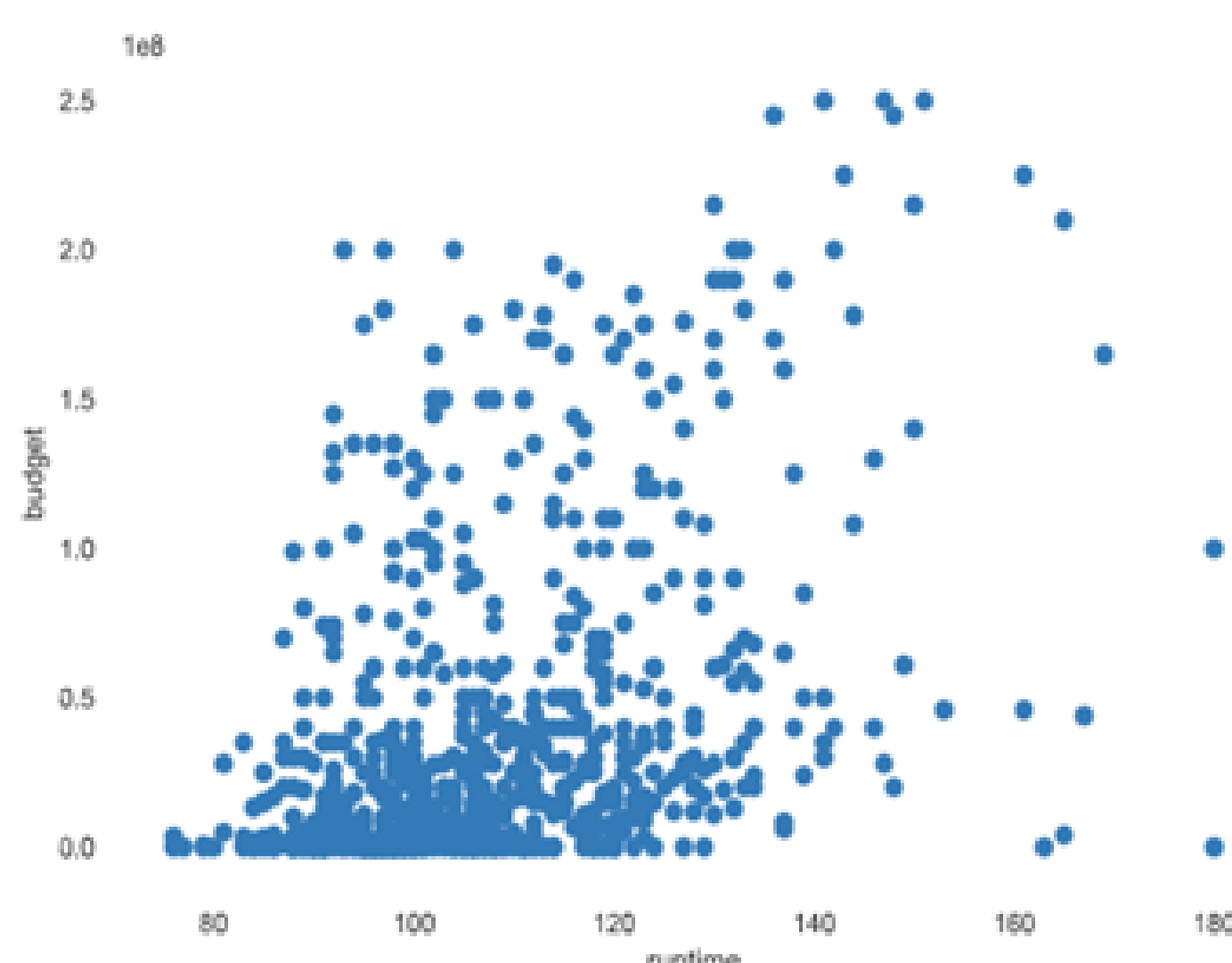
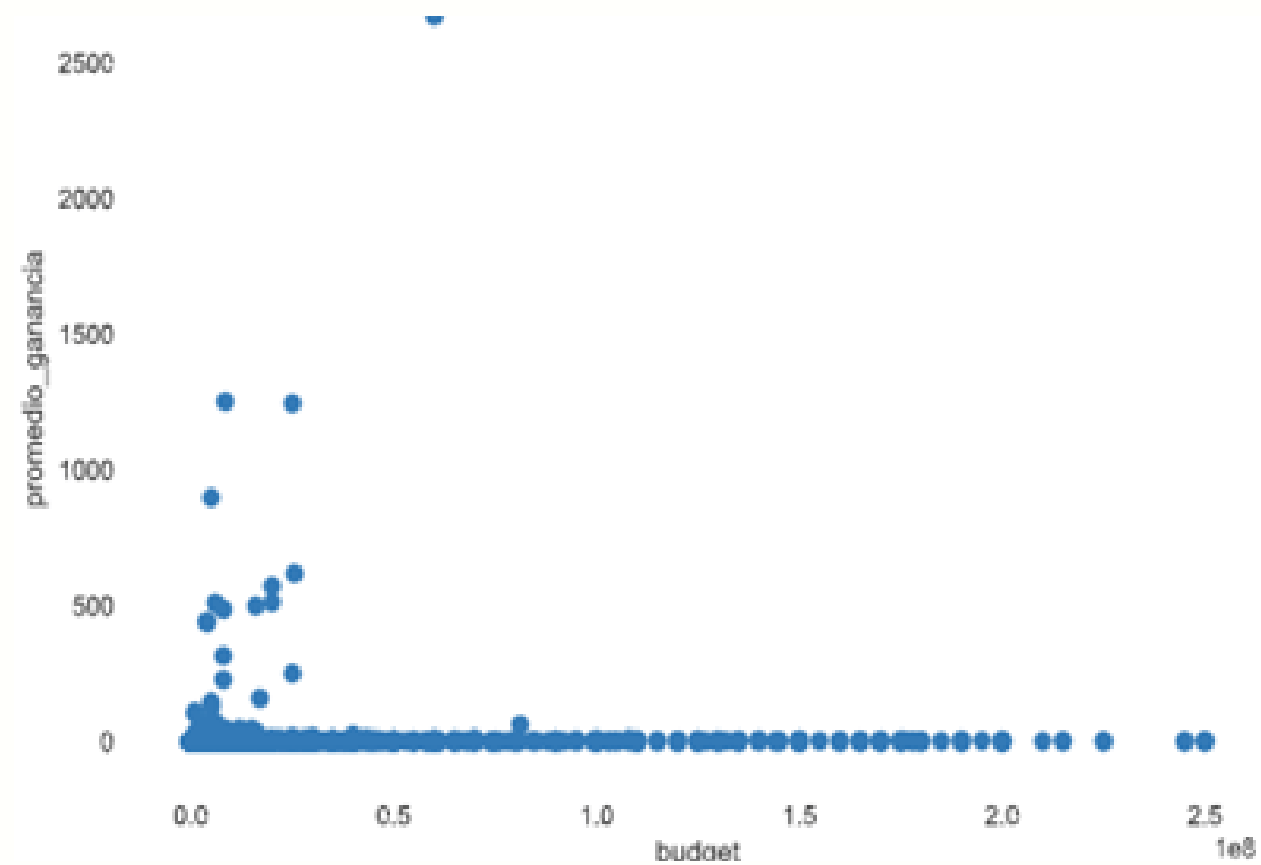
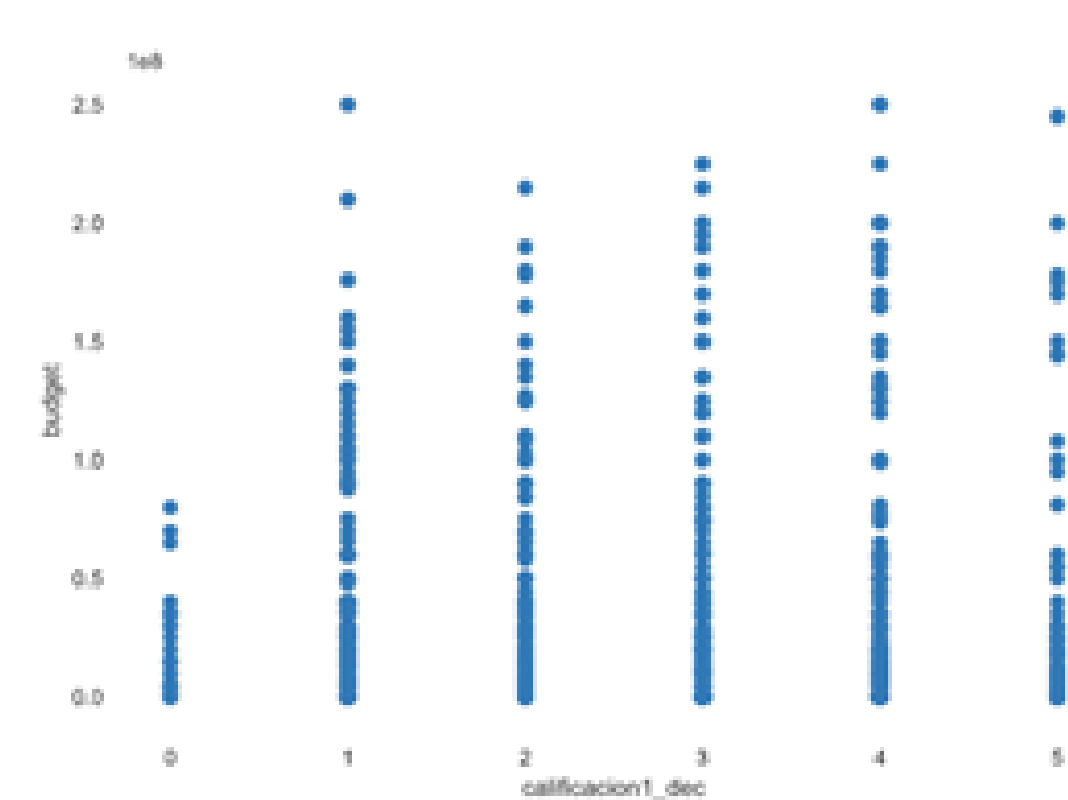
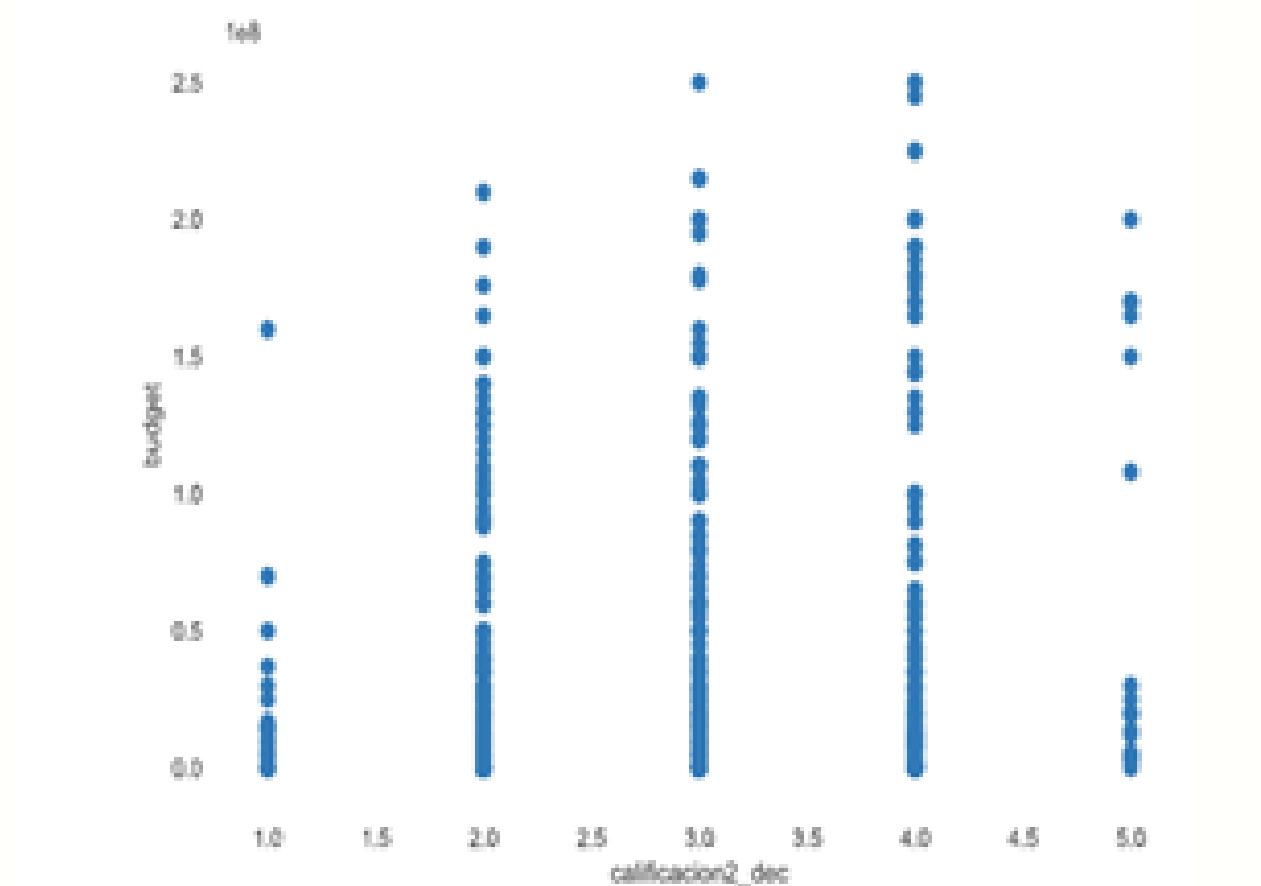
Acción, comedia y drama son los más votados

¿Duración?

La correlación entre inversión y tiempo de duración nos permite concluir que las películas de más de 120 minutos no son las mejor evaluadas.

¿Lanzamiento?

Tienen más votaciones aquellas películas que se lanzan entre octubre y diciembre



Perfila
miento
de los
datos



Acoplamiento a la estructura y fraseología
según un tipo de texto



¿Tradiciones discursivas por género?



Leyendo a Aristóteles



Leyendo a Vladimir Propp



Viendo patrones en las nubes de palabras



Viendo los resultados del K-means



estructuralismo
todo es
estructuralismo



A
interpretar
la lógica de
K-means



A interpretar la lógica de K-means



Cluster 1: la búsqueda

Historias que tienen que ver con la búsqueda: del amor, de aventura, de un objeto mágico.

Clúster 2: el don

Historias que tienen que ver con el "don": un poder, adquirir algo de la naturaleza

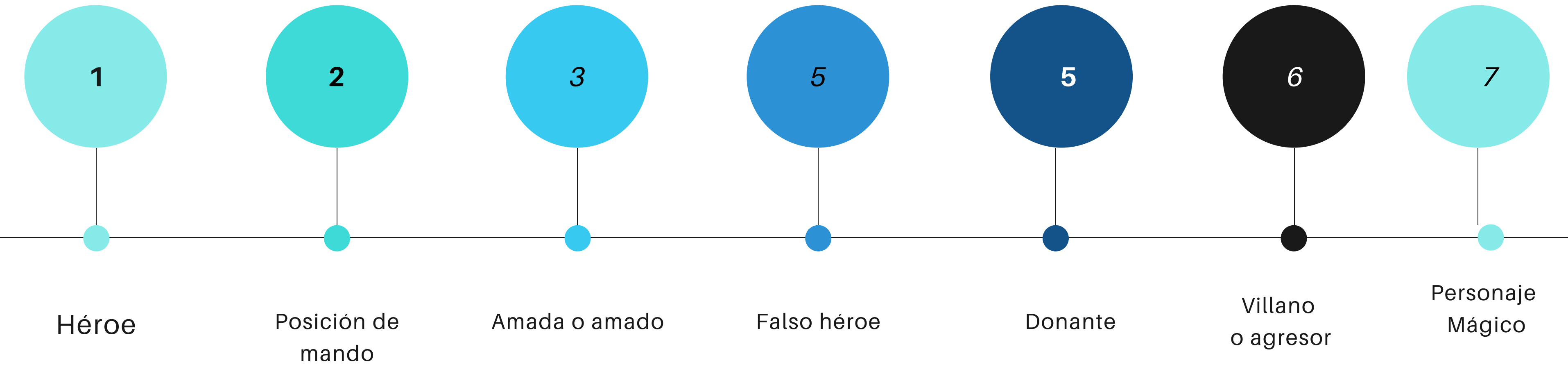
Cluster 3: no es de este planeta (ciencia ficción)

Historias de ciencia ficción (se desarrollan en un lugar extraño)

Cluster 4: encuentros

Reencontrarse o regresar a los orígenes

ROLES DE LAS HISTORIAS SEGÚN VLADIMIR PROPP



A interpretar la lógica de K-means

