UNIDAD 4: SISTEMAS DE RECOMENDACIÓN

CASO DE ESTUDIO: EL DESAFIO NETFLIX

Blanca Vázquez y Gibran Fuentes-Pineda Noviembre 2020



Desafío: construir un algoritmo de filtrado colaborativo para predecir los *ratings* de los usuarios

Contexto

- En el 2006, Netflix lanzó un desafío de un millón de dólares al equipo que lograra mejorar el rendimiento predictivo de Cinematch.
- En ese año, Netflix contaba con 7.5 millones de usuarios y 1,500 películas.
- · Problema: la matriz de ratings estaba mayormente vacía.

Cinematch

- El objetivo de *Cinematch* era rellenar toda la matriz de ratings.
- Posteriormente, generaba una recomendación, tomando las puntuaciones más altas.
- El desafío consistía llenar todas las predicciones para usuarios y películas.

Cinematch

- Cinematch tenía un RSME de 0.9524 en el conjunto de entrenamiento y un 0.9525 sobre el conjunto de prueba.
- El equipo ganador tenía que mejorar el rendimiento de *Cinematch* en un 10 %, es decir, un 0.85 sobre el conjunto de prueba.

Datos disponibles

- Netflix proporcionó un conjunto de entrenamiento de 100,480,507 ratings provenientes de 480,189 usuarios para 17,770 películas.
- Forma de los datos:
 usuario, película, fecha, rating

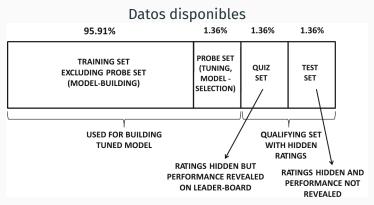


Imagen tomada de Aggarwal, 2016

Participantes

- Al inicio de la competencia 3 equipos se inscribieron, para el 2007 habían más de 20,00 equipos de más de 150 países.
- En el 2009, se presentó al ganador *BellKor's Pragmatic Chaos* los cuales habian mejorado en un 10.09 % del RSME (0.8567).



https://netflixprize.com/index.html

