

Crítica:

Collaborative Filtering for Implicit Feedback Dataset

En este paper se habla sobre un modelo que aplica un sistema recomendador basado en Collaborative Filtering, pero se distingue de los tradicionales al trabajar con datos implícitos. Hasta el momento, los sistemas recomendadores que hemos visto consisten en hacer predicciones de gustos de usuarios hacia productos en base a algún ranking entregado por el usuario explícitamente, normalmente a través de cinco estrellas. Sin embargo, la mayoría de las veces los usuarios no dan este feedback, porque les toma tiempo innecesario y no les interesa. Eso, junto al *cold start*, nos da un grave problema para poder generar buenos sistemas recomendadores. Feedback implícito no requiere que el usuario manualmente entregue información, si no que permite concluir datos a partir de comportamientos del usuario, y así contar con una base de datos mayor.

Encontré muy interesante los problemas que surgen a partir de querer clasificar un comportamiento como señal de gustos. Me hubiese gustado que el paper tratara más conversiones de comportamientos a preferencias. En particular en este problema, estaban trabajando con televisión por lo que los comportamientos pueden ser más monótonos, pero igual faltó un ejemplo de métrica útil, como podría ser lo que hace Amazon o YouTube.

Por otro lado, no entendí por qué no puede haber puntajes negativos en las estimaciones de preferencias. Comprendo que ausencia de comportamiento a favor de un ítem no indica disgusto a ese ítem, pero dado a la diferencia de disponibilidad entre dos ítems con ausencia de comportamiento, creo que debería tener mayor confiabilidad de que a un usuario no le gusta un ítem que se le ha sido ofrecido varias veces y a optado por otra cosa, que uno que nunca (o muy pocas veces) ha visto como opción.

Además, no comparto la linealidad de la métrica de preferencia en base a tiempo de visualización. En lo personal, pensaría que la visualización de un programa en televisión por un 5% de su largo es peor que 0% de visualización. Esto porque podría ser que el usuario ya vio de qué trataba el programa y decidió que no le gustó y lo cambio, versus un 0% donde simplemente aun no ha hecho un juicio (que podamos ver). Pero sí comparto que un programa visto dos veces tenga el doble de preferencia que otro solo visto una vez, dado que es un fuerte comunicador de una preferencia.

Por último, encontré de gran interés los temas abarcados. Siempre he sido escéptico de tecnologías *crowd-sourced* porque en mi experiencia personal he visto que el feedback es difícil de conseguir. Quedo con la duda de formas empíricas de traducir comportamientos a preferencias, tal vez de maneras interdisciplinarias con aporte de áreas psicológicas.

Alejandro Quiñones