

Crítica: Content-Based Recommendation Systems

En este paper analizamos los requisitos para trabajar con sistemas recomendadores basado en su contenido, tales como la representación de ítems y la representación de usuarios, para después ver distintos tipos de algoritmos usados en el rubro.

Me gustó que se realice una crítica a Tf-idf con respecto a su pérdida de contexto, pero creo que la solución entregada no profundiza lo suficiente en qué tanto mejora el contexto. Puede que hacer combinaciones de dos palabras sea muy poco, o puede que se pueda relacionar importancias de contexto a conjuntos de palabras que contienen la misma palabra para destacar su importancia dada su correlación

Con respecto a los perfiles de usuario, me pregunto si habrá análisis que permita encontrar un estadístico para la agrupación de tanto historial, donde podamos guardar toda la información obtenida en el historial, pero que la memoria utilizada no escale de tan mal manera.

La mayoría de los algoritmos ya fueron visto anteriormente, pero la utilización de métodos bayesianos es nueva y creo que es una de las mejores herramientas que tenemos para estos problemas. Me interesa saber si la fórmula implica que, después de entrenar el modelo, al recibir nuevos datos, podemos entrenar nuestro modelo, pero sin partir de cero, si no que partir la priori como la distribución obtenida post-entrenamiento. Esto haría que sistemas que reciben muchos datos con el tiempo sean más factibles

Alejandro Quiñones