

Proyecto sobre Sistemas Recomendadores Basados en Contenido

Objetivo

- El objetivo de este proyecto es implementar un sistema de recomendación basado en contenido, que nos permita recomendar los mejores documentos para un cliente, mediante el algoritmo de clasificación KNN.

Metodología

- Estudiar los modelos basados en el contenido en el material de clase.
- Crear un software que reciba un archivo de texto plano con extensión .txt, que contenga el conjunto de posibles documentos a recomendar al usuario final. Cada documento viene representado en una línea del archivo. Por ejemplo:
 1. “Aromas include tropical fruit, broom, brimstone and dried herb. The palate isn't overly expressive, offering unripened apple, citrus and dried sage alongside brisk acidity.
 2. This is ripe and fruity, a wine that is smooth while still structured. Firm tannins are filled out with juicy red berry fruits and freshened with acidity. It's already drinkable, although it will certainly be better from 2016.
 3. Tart and snappy, the flavors of lime flesh and rind dominate. Some green pineapple pokes through, with crisp acidity underscoring the flavors. The wine was all stainless-steel fermented.
 4. Pineapple rind, lemon pith and orange blossom start off the aromas. The palate is a bit more opulent, with notes of honey-drizzled guava and mango giving way to a slightly astringent, semidry finish.
 5. Much like the regular bottling from 2012, this comes across as rather rough and tannic, with rustic, earthy, herbal characteristics. Nonetheless, if you think of it as a pleasantly unfussy country wine, it's a good companion to a hearty winter stew.
 6. Blackberry and raspberry aromas show a typical Navarran whiff of green herbs and, in this case, horseradish. In the mouth, this is fairly full bodied, with tomatoes acidity. Spicy, herbal flavors complement dark plum fruit, while the finish is fresh but grabby.
 7. Here's a bright, informal red that opens with aromas of candied berry, white pepper and savory herb that carry over to the the palate. It's balanced with fresh acidity and soft tannins.”
- El software debe proporcionar como salida lo siguiente:
 - Para cada documento, tabla con las siguientes columnas:
 - Índice del término.
 - Término.
 - TF.
 - IDF.
 - TF-IDF.
 - Similaridad coseno entre cada par de documentos.



Entregables

- Enlace a repositorio GitHub público con el código fuente del sistema recomendador implementado, incluyendo un fichero README.md del repositorio, que contenga la descripción del código desarrollado y un ejemplo de uso.
- Informe en PDF describiendo el análisis realizado en varios ejemplos y las conclusiones extraídas.