



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
DEPARTAMENTO DE CIENCIA DE LA COMPUTACIÓN
IIC3633 - SISTEMAS RECOMENDADORES

Informe Intermedio

15 de noviembre de 2019

2º semestre 2019 - Profesor: Denis Parra

Ariel Martínez y Jerónimo Salazar

Propuesta original

En la primera entrega del proyecto (llamada propuesta) presentamos la posibilidad de implementar recomendaciones basadas en contenido haciendo uso de las métricas TF-IDF. Además, propusimos trabajar con un *dataset* de transcripciones a texto de charlas TED. Sin embargo, dado el *feedback* de los ayudantes y el profesor, nos convencimos de que la mejor idea sería pivotear hacia otra técnica de recomendación y otro dominio por las siguientes razones (sin ningún orden en particular):

- TF-IDF es comúnmente utilizado en sistemas recomendadores, sin embargo, no es considerado un algoritmo de recomendación en sí, por consiguiente, la propuesta entregada era insuficiente como proyecto de este curso.
- TF-IDF y todas sus variantes y técnicas relacionadas ya han sido exhaustivamente estudiadas, dando poco o nulo espacio de exploración para el desarrollo del presente trabajo.
- El *dataset* a utilizar no contenía interacciones de usuario, sino que únicamente agregaciones de ellos y *rankings* de origen desconocido, por lo que no hubiese sido posible evaluar las recomendaciones generadas por nosotros de manera fidedigna.

Nueva propuesta

El tema escogido para continuar con el proyecto es la **recomendación multi criterio**, un tipo de filtrado colaborativo muy bien explicado en el capítulo Multi-Criteria Recommender Systems del Recommender Systems Handbook. Acá, Adomavicius et. al. (2011) explican que este tipo de recomendación puede ser expresado como un problema de decisión, que requerirá realizar cuatro pasos de manera preliminar (en este orden):

1. Definir el **objeto de decisión**. Esto es, escoger el conjunto de alternativas sobre las

- cuales se debe tomar la decisión (i.e. los ítems). Aquí, identifican cuatro problemáticas diferentes: *choice*, *sorting*, *ranking* y *description*, cada una enfocada en trabajar sobre una categoría de objeto en particular.
2. Determinar la **familia de criterios**. Es decir, identificar y especificar la o las funciones que modelen las preferencias de quien toma la decisión (i.e. el usuario). En este caso, especifican cuatro tipos de criterio: *measurable*, *ordinal*, *fuzzy* y *probabilistic*.
 3. Desarrollar un **modelo de preferencia global**. Esto hace referencia a sintetizar las preferencias parciales por cada criterio y agregar dichos valores en una sola función de utilidad escalar. Aquí, describen cuatro grupos de modelos: *value-focused*, *optimization*, *outranking relations* y *other*.
 4. Seleccionar el **soporte del proceso de decisión**. Esto es, diseñar el sistema, los métodos y el procedimiento a través de los cuales se llevará a cabo la toma de decisiones.

En este mismo trabajo, Adomavicius et. al. enumeran una serie *papers* divididos de acuerdo a las clasificaciones mencionadas en los puntos 1, 2 y 3 de la lista anterior.

En el presente proyecto, nuestro plan a seguir consiste en implementar un algoritmo de recomendación multi criterio (ocupando los distintos aspectos de las cervezas que nuestro *dataset* posee) y comparar su rendimiento con una recomendación con criterio único (ocupando solamente la calificación global de cada ítem). Este algoritmo lo buscaremos entre los *papers* enumerados por Adomavicius et. al., de manera tal que se adecúe el *set* de datos con el que trabajaremos y compararemos el funcionamiento de una recomendación de filtrado colaborativo con respecto al algoritmo multi criterio seleccionado, como también el funcionamiento de este mismo algoritmo, normalizando los valores de las distintas calificaciones ocupadas.

Trabajo futuro y datos

Los datos con los que proponemos trabajar son 1.586.614 *reviews* (comúnmente traducido como críticas o reseñas) de cervezas pertenecientes a un lapso de 10 años (entre 2001 y 2011), recopiladas por [Beeradvocate](#) — un sitio *web* dedicado al *rating* de cervezas creado en 1996. Esta base de datos se encuentra originalmente en formato CSV, donde cada fila corresponde a la interacción de un usuario con una cerveza (un ítem consumido por él).

Entre los atributos de la tabla es importante destacar el identificador único de cada ítem (`beer.beerid`), junto con otros elementos relacionados a éste (como el nombre, el estilo, la cervecería y la concentración de alcohol) y el nombre del usuario calificador (`review.profilename`). Además, la columna `review.time` revela el tiempo en el que se registró cada entrada a través de un *timestamp*. Entonces, con referencia a los pasos preliminares mencionados en la sección anterior, sería correcto determinar que el objeto de decisión es o son las cervezas. Dado esto, nosotros estamos interesados en realizar un *ranking* de cervezas para un usuario arbitrario.

Por otro lado, los criterios a través de los cuales un usuario califica una cerveza son 5 y todos numéricos, con valores entre 1 y 5 (`review.overall`, `review.aroma`, `review.appearance`,

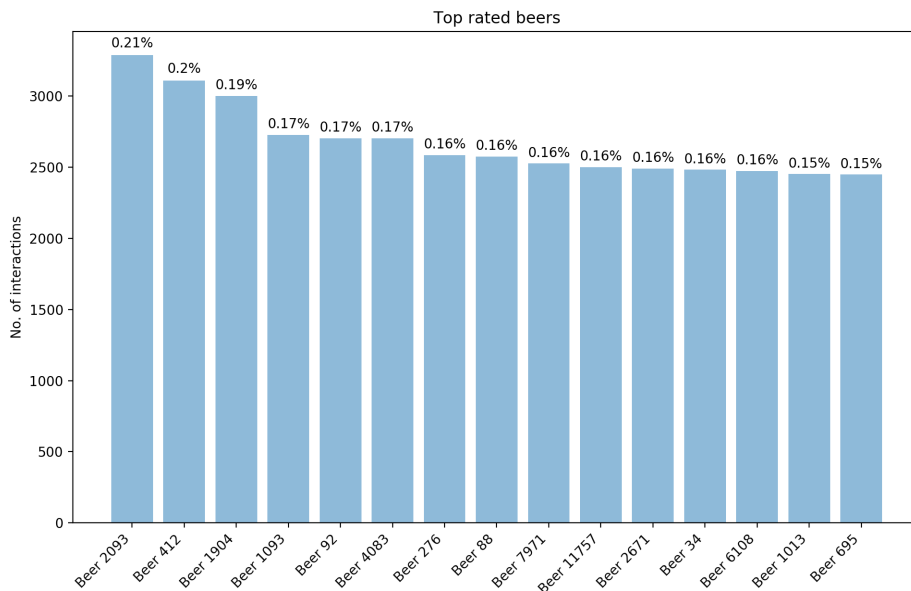


Figura 1: Top 15 cervezas con mayor número de calificaciones

`review_palate` y `review_taste`), por lo que pueden incluirse en su totalidad en la familia de criterios *measurable*. La ventaja que tiene esta familia es que pueden crearse métricas de distancia de manera intuitiva, facilitando el proceso de comparación de ítems. En este contexto, Adomavicius & Kwon (2007) proponen varias técnicas interesantes para realizar *ranking* multi criterio, haciendo uso de variables cuantitativas y funciones de preferencia global *value-based*.

En el presente proyecto, nuestro plan a seguir consiste en implementar un algoritmo de recomendación multi criterio (ocupando los distintos aspectos de las cervezas que nuestro *dataset* posee) y comparar su rendimiento con una recomendación con criterio único (ocupando solamente la calificación global de cada ítem). Este algoritmo lo buscaremos entre los *papers* enumerados por Adomavicius et. al., de manera tal que se adecúe el *set* de datos con el que trabajaremos y compararemos el funcionamiento de una recomendación de filtrado colaborativo con respecto al algoritmo multi criterio seleccionado, como también el funcionamiento de este mismo algoritmo, normalizando los valores de las distintas calificaciones ocupadas.

Para comparar el rendimiento de los distintos métodos, se ocuparán las métricas MAP y nDCG. Separaremos el dataset, dejando un 90 % de las calificaciones como dataset de entrenamiento, mientras que el restante 10 % será utilizado como dataset de prueba para la recomendación y cálculo de las métricas recién mencionadas

Número total de calificaciones	1.586.614
Número total de usuarios	33.388
Número total de items	66.055
Promedio de número de calificaciones por item	24,02
Complejidad de matriz usuarios/items	0.001

Cuadro 1: Caption

Promedio	
Overall	3.816
Aroma	3.736
Apariencia	3.842
Paladar	3.744
Sabor	3.793
Promedio características	3.778
Desviación Estándar	
Overall	0.721
Aroma	0.698
Apariencia	0.616
Paladar	0.682
Sabor	0.732
Promedio características	0.682

Cuadro 2: Métricas de cada tipo de calificación

Exploración de datos

Realizamos un análisis a los datos obtenidos de *Beeradvocate*. De este análisis pudimos obtener las métricas e información mostradas en la Figura 1 y Tabla 1. Como es de esperar, la matriz usuario/cerveza es poco completa (tiene una completitud de menos de 1 %) y existe una gran diferencia entre el promedio de calificaciones por item, contra los items con mayor calificación (aproximadamente 3000, comparado con el promedio de calificaciones que es 24).

Por otra parte, en la Tabla 2 se muestran las métricas de cada calificación. De esta podemos darnos cuenta que los promedios y desviaciones estándar se encuentran en rangos cercanos y que tanto el promedio como la desviación estándar del *overall_review* tienen una diferencia porcentual de, aproximadamente, 1 % y 5,4 % respectivamente, con los valores obtenidos al promediar los promedios y desviaciones de los demás aspectos que se califican (aroma, sabor, paladar y apariencia).

Referencias

1. Adomavicius G. & Kwon Y. (2007). *New Recommendation Techniques for Multi-Criteria Rating Systems*. Intelligent Systems, IEEE. DOI: 10.1109/MIS.2007.58.
2. Adomavicius G., Manouselis N. & Kwon Y. (2011). *Multi-Criteria Recommender Systems*. Visto en: [3]. Disponible en:
https://link.springer.com/chapter/10.1007/978-0-387-85820-3_24.
3. Ricci F., Rokach L., Shapira B. & Kantor P. B. (2011). *Recommender Systems Handbook* (capítulo 24). Springer, Boston, MA. ISBN: 978-0-387-85820-3.
4. *Beer Reviews from the Beeradvocate* (2011). TCB Analytics. Disponible en:
<https://data.world/socialmediadata/beeradvocate>.