

Pattern Name: *Content-based Recommendation*

Aliases: *Content-based filtering*

Context

In a domain-specific database, when users search for a subject, other similar items based on the subject content are to be recommended, thus the user can consume more items over that database. The recommendation should be made using the subject's characteristics or its metadata. The consummation usually is related to a profit for the provider (over the subject itself, like on movie rental, or over some advertising attached to subject, like on news or social media webpages), but it can also be related only for the system popularity growth.

Problem

How to recommend other similar items when a user searches for a specific subject?

Forces

Forces influencing the problem and solution are:

- High accuracy related items rises chances to the user effect consume more items of his wish that haven't been considered yet
- User profile does not have so much influence on the recommendation
- User can rate items that he already have evaluated

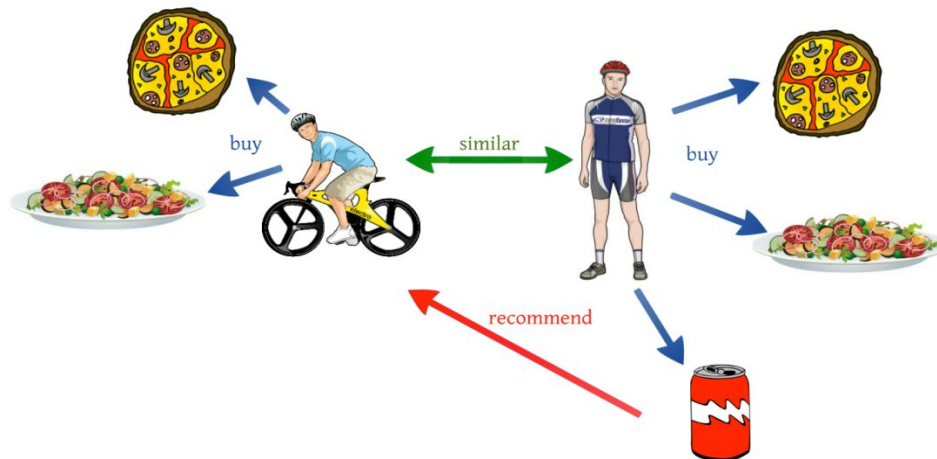
Solution

Apply a recommendation algorithm that emphasizes more on items attributes analysis in order to generate predictions. Items that are mostly related to the positively rated items are recommended to the user. The algorithm can use different types of models to find similarity between items in order to generate meaningful recommendations. These models make recommendations by learning the underlying model with either statistical analysis or machine learning techniques. Also, if the user profile changes, the algorithm still has the potential to adjust its recommendations within a very short span of time.

Known Uses

Citeseer is a research paper repository that uses the content-based recommendation. It utilizes heuristics and machine learning algorithms to process documents and index its content. Movies recommendation is also an example that the algorithm can be made based on its attributes (metadata), by showing a list of movies with the same genre, cast, director, etc. eg.: IMDb, See this Next.

Sketch



Withdraws

This strategy of recommendation depends on items' metadata and how much their characteristics are refined. Besides, the algorithm can become over specialized as the system starts to recommend only similar items already defined in the user profile.

Related Patterns

Content-based recommendation can use several types of models to find similarity, like Vector Space Model, Term Frequency Inverse Document Frequency (TF/IDF). It can also use probabilistic models such as Naïve Bayes Classifier, Decision Trees or Neural Networks to model relationship between different documents within a corpus, either by statistical analysis or machine learning techniques. [1]

Another type of recommendation made over domain-independent subjects is known as Collaborative Filtering. It consists in building a database based on user's items preferences, so different types of subjects can be crossed in a user-item matrix. Thus, the Collaborative technique makes possible recommendations like "who bought this also bought that".

Author: Luiz Wagner Tavares Nascimento

nascimento1wtn@gmail.com

Date: 04/18/2016

Pattern Source: INPE CAP-392 Design Patterns 2016

References

[1] Park, Deuk Hee, et al. "A literature review and classification of recommender systems research." *Expert Systems with Applications* 39.11 (2012): 10059-10072.