Recommender System based on publicly available data

Siddhi Deshpande, Rajani Karbhari, Nikita Khandekar, Sharayu Suryawanshi, Prof. Pranoti D.Kale

Abstract: Recommender systems are meant for recommending products to customers according to their interests. Recommender system has several features namely, data collection and processing, recommender model, recommendation post processing and user interface etc. in order to recommend proper products to given user. These recommendations systems rely on one or more recommendation techniques. The paper proposes a recommender system that will recommend the products that are relevant to user's interests in different fields/areas/domains. User interests are extracted with the help of his/her activities in social networking sites such as Facebook. Generally, recommender systems are recommended the products or services regarding one specific domain whereas the proposed system is able to recommend the products from various domains. Along with that, system can recommend text articles as well that user might find on his/her interests

Keywords: web crawling algorithms, crawling algorithm survey, search algorithms

Introduction:

There is a large amount of E-commerce websites with various products from many different categories. As a result, there's abundant data or number of products everywhere for user to search through. Although this provides wide range of products to surf through for user, user can find most of the products irrelevant to his/her interest. That's why recommender systems came into picture. Recommender systems recommend product in accordance with user's interests. These interests are found with the help of the products that user surfs through or checks regularly. Most recommender systems rely on collaborative filtering which groups the customers having similar interests and recommends the products that users find of their interests to each other. Apart from that, social networking sites can provide the information regarding interests of user. Thus, paper proposes a system that can identify the interests of user from his/her social networking profiles and search products with the help of recommender systems which collects information about products through RSS feeds, different APIs and web crawling systems etc. Using these two information, system then recommends products that can be more relevant to user's interests. Using a social media information to feed recommender system with the data is present in some previous papers. Some authors calculate similarities between social network users based on data obtained from the Facebook API. They used acquired data to recommend a tourist attraction to users they consider to have similar taste and some of the author also used data extracted from a real online social network to overcome the data sparsity and the cold-start issues. Recommender systems for the movie and music domain are the focus of many research papers. Some dealing with a movies recommendation are based on movie genre correlations while others generate recommendations using various types of data - about users, the available items, and previous transactions. Our recommender system leverages the existing publicly available recommendations APIs for the movie and music domain to get an initial set of recommendations for the given genre or author. We further enrich that data based on users feedback and thereby improve subsequent recommendations.

Motivation:

Recommender system is meant for recommending set of products to user that user can find relevant to his/her interests. Apart from the products that a user is surfing through, there are lots of other ways to find user's interests. Information from his/her social networking profile is one of them. Users tend to post things that can tell about their areas of interests on social networking sites like Facebook or twitter. This information should be put into a better use when recommending products to that user. A system that can identify the user's interest from social activities of user should be built in order to do so.

System Architecture:

Proposed system uses social networking as a source of data and according to user’s interest system will recommend link by using crawler algorithm.
Crawler Algorithm:

A crawler is a program that visits Web sites and reads their pages and other information in order to create entries for a search engine index. The major search engines on the Web all have such a program, which is also known as a “spider” or a “bot.” Crawlers are typically programmed to visit sites that have been submitted by their owners as new or updated. Entire sites or specific pages can be selectively visited and indexed. Crawlers apparently gained the name because they crawl through a site a page at a time, following the links to other pages on the site until all pages have been read.

Conclusion:

Recommender systems are a powerful technology for the entertainment industry and social networks. We presented recommender system that exploits existing publicly available services to gather data needed to build the user profile and to generate initial recommendations. User’s interests are deduced from his activities and posts in a social network. Using this data for building user profile we bridge the cold start problem. Currently, we have implemented recommendations for URL-link related. We connected concepts from user profile with external items obtained through Open API calls. These and similar items are subsequently recommended to the user. To get feedback from users we implemented recommendations grading. Those way users affect future recommendations because grades change and tune user profiles. Recommender system evaluation showed that generated recommendations.

Reference:


