www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 3, Issue 10 October, 2014 Page No. 8988-8992

Leveraging Hybrid Recommendation System In Insurance Domain

Sanghamitra Mitra, Nilendra Chaudhari , Bipin Patwardhan

Technical Architect
Research & Innovation, iGATE Global Solutions, Mumbai, India
Sanghamitra.Mitra@igate.com

Technical Architect
Research & Innovation, iGATE Global Solutions, Mumbai, India
nilendra.chaudhari@igate.com

Senior Technical Architect
Research & Innovation, iGATE Global Solutions, Mumbai, India
bipin.patwardhan@igate.com

Abstract: ***Recommendation algorithms are best known for their use on e-commerce web sites, where they are used to create additional business opportunities by suggesting additional products and services. Generally, the recommendation are created by collating feedback from various users who have purchased the same or different products, as well as comparing the features of the products themselves. While recommendations have been most successful in domains like retail, due the availability of large volume of feedback, it is challenging to implement in domains where there is no such prior information or the information available is very small in volume. This paper presents how a hybrid recommender system was leveraged in insurance domain, by integrating an attribute based recommendation system with preference-based recommendation system.***

Keywords: Recommendation, Insurance, Collaborative Filtering, Attribute-based Recommendation.

1 Introduction

Recommender Systems (RSs) are software tools and techniques that can be used to provide suggestions for items. Suggestions can relate to various decision-making processes, such as what items to buy, what music to listen to, or what online news to read.

'Item' is the general term used to denote what the system recommends to users. An RS normally focuses on a specific type of product or service, its design, its graphical user interface, as well as the recommendation technique. All of these put together, make for an effective platform that can serve useful suggestions for the products being perused by customers and improve the chances of the customer choosing one of the recommended products or services.

The development of an RS was from a rather simple observation: individuals often rely on recommendations provided by others while making routine, daily decisions. For example, it is common to rely on what one’s peers recommend when selecting a book to read; employers count on recommendation letters in their recruiting decisions; and when selecting a movie to watch, individuals tend to read and rely on the movie reviews that a film critic has written and which appear in the newspaper they read.

While it is not possible to have personalization in media like newspapers, TV and radio, most users depend on their social circle to get feedback regarding the item they wish to purchase. This personalization is now possible on e-commerce and web sites, where user specific suggestions can be served/displayed by analyzing their previous interaction with the website and if available, the feedback from their social circle.

While recommendations can take multiple forms, in their simplest form, recommendations (personalized or otherwise) are offered as ranked lists of items. Due to personalization, different users get to see different recommendations which improve the chances of conversion into a concrete sale, as user preferences are taken into consideration.

1.1 Use of Recommendation System

In the previous section we defined RS for providing users with suggestions for items. We would like to refine this definition, illustrating a range of possible roles that a RS can play. First of all, we need to distinguish between the roles played by the RS on behalf of the service provider, from that of the user of the RS.

Service providers typically wish to use this technology to increase the number of items sold, sell more diverse items or to increase the user satisfaction.

In comparison, the user's primary motivation is to find a suitable products or service which will cater to their specific need or which they are not presently aware of.

2 Types of Recommendation Systems

Nowadays, RS are used in a wide variety of applications. Each of these applications has its particular characteristics, with greatly differing temporal dynamics or volatility, amount of available data, use of explicit or implicit indicators, etc. At the same time, there are various approaches and techniques to build a recommendation system. The appropriate techniques are to be selected which would cater to the current business need

The most common models required to introduce an RS are

- Cluster-based Model,
- Preference-based Model,
- Attribute-based Model, and
- Hybrid Model

2.1 Cluster-based Model

In this model, clustering techniques are used to identify similar products and services. To find customers who are similar to the user, cluster models divide the customer base into a set of segments. Segments are typically created using a clustering or a similar learning algorithm, although some applications use manually determined segments.

Once the algorithm generates the segments, it computes the user's similarity to vectors that summarize each segment, chooses the segment with the strongest similarity and classifies the user accordingly. Some algorithms classify users into multiple segments and describe the strength of each relationship.

Cluster models have good online scalability and performance as they compare the user to a controlled number of segments rather than the entire customer base. However the recommendation quality is low as the similar customers that the cluster models find may not the most similar customers always, hence the recommendations they produce are less relevant.

2.2 Preference-based Model

Currently, most RSs are based on user preference or user feedback which uses collaborative filtering algorithms. Collaborative filtering algorithm represents a customer as an N-dimensional vector of items, where N is the number of distinct catalog items. The components of the vector are positive for purchased or positively rated items and negative for negatively rated items.

To compensate for best-selling items, the algorithm typically multiplies the vector components by the inverse frequency (the inverse of the number of customers who have purchased or rated the item), recommending less well known items to the users. The algorithm generates recommendations based on a few customers who are most similar to the user in terms of their preference to the items.

It can measure the similarity of two customers, A and B, in various ways, based on various similarity metrics such as cosine measure, Pearson correlation, Euclidean distance, Tanimoto coefficient, etc.

Even though preference based approach provides good result in many cases, it has major downsides when dealing with new users or items that are not sufficiently connected in the known data (cold start problem).

2.3 Attribute-based Model

To overcome the challenges of preference based recommendation system, attribute-based model analyses the data related to user or the item. The basic process performed by an attribute-based recommender consists of matching the attributes of a user profile in which preferences and interests are stored, with the attributes of a content object (item), in order to recommend new and/or interesting items.

The system implementing the attribute-based model that analyses the content consists of information or attributes about the item or the user. In this approach a similarity matrix is computed based on the user or item attribute. This matrix forms the basis for the recommendations.

This approach improves recommendation quality by using addition data content - especially for new users and new offers - by providing a solution to the cold start problem. However due to lack of readymade implementation, the development cost would be on the higher side.

2.4 Hybrid Recommendation

Both, the preference-based model and the attribute-based model have advantages and limitations. Hence the most common approach is to use hybrid recommendation, by combining attribute-based recommendations with preference-based recommendations.

However, hybrid approaches most commonly integrate information about the recommended items, which can be modeled as a similarity metric between items and thus integrates easily in an item-based recommender model. Attribute-based models are therefore useful to deal with cold start for new content, but do not help with recommendations for new users.

Other recommenders use user neighborhoods (user based recommendation) which are easier for integrating demographic information, but few systems can include both user and item information in combination with interaction data.

3 Related work

In recent years, the interest in recommender systems has dramatically increased, as the following facts indicate:

1. Recommender systems play an important role in Internet sites such as Amazon.com, YouTube, Netflix, Yahoo, Tripadvisor, Last.fm, and IMDb. Moreover, many media companies are now developing and deploying RSs as part of

services they provide to their subscribers. For example Netflix, the online movie rental service, awarded a million Dollar prize to the team that first succeeded in substantially improving the performance of its recommender system.

2. There are dedicated conferences and workshops related to the field. We refer specifically to ACM Recommender Systems (RecSys), established in 2007 and now the premier annual event in recommender technology research and applications. In addition, sessions dedicated to RSs are frequently included in the more traditional conferences in the area of data bases, information systems and adaptive systems. Among these the following are worth mentioning, namely ACM SIGIR Special Interest Group on Information Retrieval (SIGIR), User Modeling, Adaptation and Personalization (UMAP), and ACM's Special Interest Group on Management Of Data (SIGMOD).

4 Recommendation Pilot

Though RS are most commonly used in the retail domain to provide suggestions for related products they are not so widely used in other domains like Insurance. We believe that the Insurance industry can be benefited by using an RS. This was the motivation for exploring this area and executing a pilot with one of our customers. Our aim was to extend recommendation concept for insurance industry and demonstrate the usefulness of preference-based and attribute-based model for creating cross-selling and up-selling opportunities for other products and services.

The pilot solution covers recommendation of insurance products and riders associated with insurance products purchased by the user. Here we have used hybrid model of RS mentioned in above section. Figure 1 depicts how customer satisfaction can be improved by providing personalized recommendation for insurance products and riders through various channels like customer care, agents, websites and others

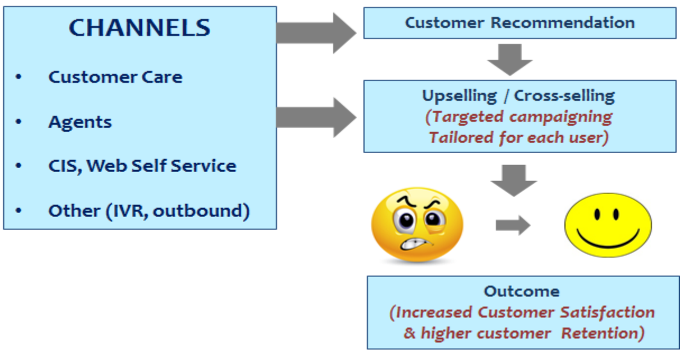

Figure 1 : Business Use Case

4.1 Policy Recommendation

Insurance is important part of financial planning and there are different attributes which trigger (significant change in the lifestyle) the need to get insured like marriage, birth or adoption of child, buying home/second home, starting business, retirement, buying vehicle etc. It has been observed that a majority of people do not have insurance that cater to such triggers. Presently, insurance companies are not able to benefit from such triggers due to various reasons like non reliable agents, lack of knowledge amongst customer support staff, user think nothing will happen to them etc. By using suitable RS, it is possible to cater to such triggers and the Insurance Company can suggest suitable products to their customers.

To provide an effective and useful insurance policy that can address the triggering event, a policy can be suggested by identifying those users who are similar to the user being considered.

For the pilot, we have implemented policy suggestions using attribute-based recommendation as shown

Figure 2.Step for recommendations are creating user attribute matrix, computing similarity amongst user, finding top n similar user (Adam Daniel David) and recommending policies based on similarity matrix.

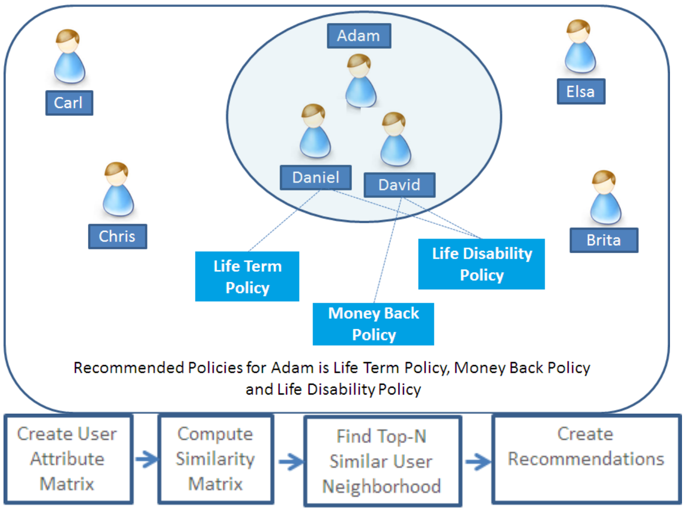

Figure 2 : Introduction to Attribute Based Recommendation

4.2 Rider Recommendations:

Riders are provisions of an insurance policy that is purchased separately from the basic policy and that provides additional benefits (at additional cost). For example, the critical illness rider covers a customer in the event that she is diagnosed with any of the critical illness like heart attack, cancer, stroke, etc. By purchasing this rider, the insurer is paid lump sum amount, to take care of the critical illness.

Riders are important, as they can be used to customize an insurance policy to the specific needs of the customer. Given the technicality of the policies, customers usually find it difficult to choose riders while purchasing an insurance policy. As riders have to be purchased along with the policy itself, choosing the riders is a difficult task. Unlike the insurance policy, in most cases, the decision to purchase riders is based on user behavior/taste more than any trigger. E.g. customers opting for critical illness rider will be decided based on user behavior rather than any other attribute. To make the task of choosing a rider easy, the pilot implementation uses an RS to suggest riders, by collating feedback from other customers. As the recommendations consider feedback from other users, the chances of the customer purchasing the rider, are increased. The pilot implementation uses preference based model to identify suitable riders.

As shown in Figure 3 user rates various riders which they use. Comparison of rating shows that user5 is most similar to user1 followed by user4. Riders highly rated by user4 and user5 will be most likely be recommended for user1.

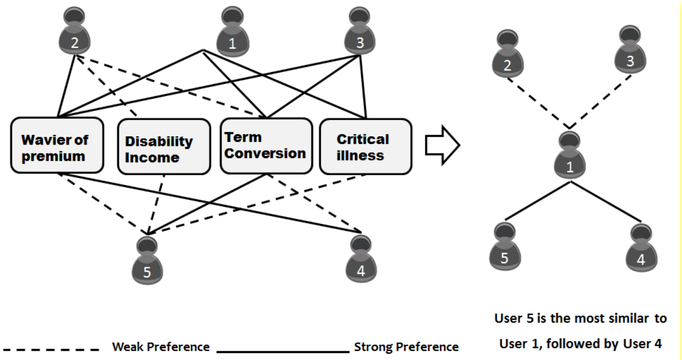

Figure 3 : Collaborative Filtering for Riders

4.3 Solution Architecture

Figure 4 depicts the high level architecture used to implement the pilot. The architecture is scalable as it is built using Big Data platforms - Apache Mahout and Apache Hadoop.

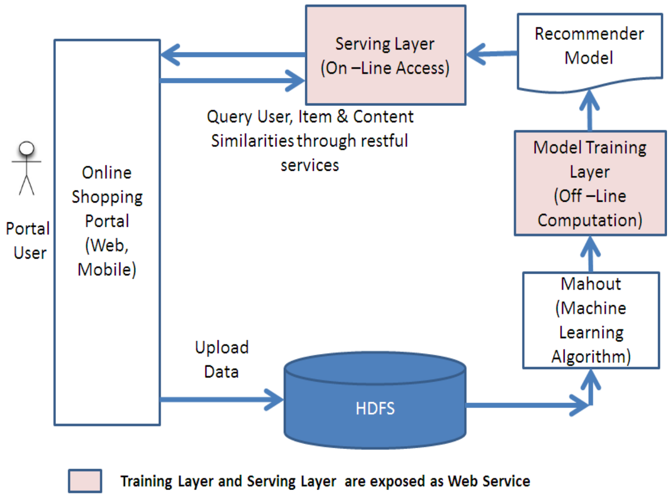

Figure 4 : High Level Architecture

The architecture is broadly divided into three major components namely, Model Training layer, Serving layer and Portal.

- **Model Training Layer:** This layer is used to build the recommender model which can be used by the Serving layer for providing recommendation. This layer requires rating data, user data and item data, in order to create the model.

- **Serving Layer:** This layer uses the model built in the training layer to generate recommendation based on the request coming from the Portal. Here, users can get recommendations based on collaborative filtering or content-based recommendation.

- **Portal:** This is user interface for policyholders where users can view the existing policies and purchase new policies. Recommendations for policies are riders are integrated in the portal.

5 Benefits

The solution helps the insurance companies in various ways to provide personalized service in a cost effective manner. Some of the benefits of this solution are:

- **Personalized Cross-selling/Up-selling:** The system helps customers to discover new and not-yet-experienced services that would be relevant to their needs which leads to better customer satisfaction.

- **Reduced Total Cost of Ownership:** Data driven recommendations reduce the dependency on highly trained sales and customer support teams as well as agents to recommend appropriate service, thereby reducing the total cost of ownership of marketing.

- **Increased Reach:** The hybrid recommendation approach will be able to cater to wide variety of business use cases.

- **Resolves Cold Start Problem:** As the hybrid approach of recommendation considers the user or item attributes for recommendation along with the user preference, it solves the cold start problem posed by the conventional preference based recommendation.

- **Scalable:** Preference-based recommendation needs to deal with sparse matrix which may not be a scalable solution in all cases. Suitable selection of an algorithm and input data based on the business use case makes to solution more scalable.

- **High Performing:** Suitable use of big data technologies including Mahout, Hadoop and NOSQL technologies makes the solution high performing.

6 Conclusion and Future Work

Preference-based model depends upon user preference data for generating recommendations. However, in the insurance domain we have observed that user preference information is inadequate. To overcome this limitation, we have implemented a hybrid model wherein attribute based recommendation is combined with preference based recommendation to generate better quality recommendations.

While the hybrid approach of recommendation systems is a good fit for the insurance domain, based on specific business need, the recommendation system can be extended. The current preference and content attribute can be augmented with other set of context dataset such as social data, location information while context mining may provide more accurate result.

Along with collaborative filtering and similarity based algorithms other algorithms such as graph based approach can also be leveraged to cater to a specific business need.

References <SM>

[1] Leo Iaquinta, Anna Lisa Gentile, Pasquale Lops, Marco de Gemmis and Giovanni Semeraro "A Hybrid Content-Collaborative Recommender System Integrated into an Electronic Performance Support System", Seventh International Conference on Hybrid Intelligent Systems, 2007.
[2] Guangping Zhuo, Jingyu Sun and Xueli Yu "A Framework for Multi-Type Recommendations", Eighth International Conference on Fuzzy Systems and Knowledge Discovery, 2007.
[3] Yechun Jiang, Jianxun Liu, Mingdong Tang and Xiaoqing (Frank) Liu "An Effective Web Service Recommendation Method based on Personalized Collaborative Filtering", 2011 IEEE International Conference on Web Services.
[4] Chuangguang Huang and Jian Yin "Effective Association Clusters Filtering to Cold-Start Recommendations", 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery.
[5] Lior Rokach and Guy Shani and Bracha Shapira and Eyal Chapnik and Gali Siboni "Recommending Insurance Riders"
[6] Mustansar Ali Ghazanfar and Adam Prugel-Bennett," A Scalable, Accurate Hybrid Recommender System", 2010 Third International Conference on Knowledge Discovery and Data Mining.
[7] Liang He and Faqing Wu, "A Time-context-based Collaborative Filtering Algorithm"
[8] Qian Wang, Xianhu Yuan, Min Sun "Collaborative Filtering Recommendation Algorithm based on Hybrid User Model", FSKD, 2010.
[9] Jong Seo Lee "Survey of Recommender Systems (Collaborative Filtering)"
[10] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: item-to-item collaborative filtering," IEEE Internet Computing, vol. 7, no. 1, pp. 76–80, 2003.

Author Profile

Sanghamitra Mitra got her diploma in computer application from Indian Statistical Institute and working as a Technical Architect with more than 15 years of experience in the IT industry. She has experience of working with international clients including British Airways, Hitachi Medical Corporation, Car Phone Warehouse. At IGATE Global Solutions, she works with Research and Innovation group and responsible for exploring and incubating emerging technologies.

Nilendra Chaudhari is a Technical Architect having 13 years of experience in end to end system development. He has worked across different domains for clients such as eBay, GE, Symantec, J&J and State Bank of India associates. He has technical expertise on wide range of technologies like Visual Basic, Java, J2EE, Adobe Flex, Big Data, In-memory Computing and Parallel Programming.

Bipin Patwardhan is a Sr. Technical Architect with more than 17 years of experience in the IT industry. At IGATE Global Solutions, he is part of the Future and Emerging Technologies group. The group explores new and emerging technologies including 3D printing, wearable computing, analytics and augmented reality. He was earlier leading the High Performance Computing CoE and the Big Data CoE in the group