

Data Mining Business Opportunity Write-up

- Pandu Ranga Madasu

Firm Selection and Background:

The firm considered for this write-up is **Amazon.com, Inc.** It is the largest Internet based retailer in the world by total sales and market capitalization. From food to apparel, books to electronic goods it sells diverse range of products and has become an undisputed leader in the e-commerce industry. What's unique and interesting about Amazon is its widely popular Recommendation models, which have been proven successful over the years. It recommends a range of products from different categories to each user based on their searching patterns, with the aim of suggesting items to the users that they are more likely to click, learn and ultimately buy.

Business Opportunity for data mining:

The focus here is on the existing customers that transfer to a new country or to a new place due to various reasons such as job transfers, pursuing higher education or vacations. And the aim is to build a model that learns from the historical purchase data of the similar kind of customers in the past and recommend the current customers a set of products that they are more likely to buy once they move-in such as bedding, Kitchenware and any other items which are difficult to carry or are not allowed during the travel. All these products discovered from the purchase patterns can be grouped into a new product category that can be named as 'Move-in kit'. Every year, there will be hundreds of thousands of customers in either of these categories shifting to a new place. International students alone make up to nearly one million population every year. And since there will be baggage size limitations and safety measures, certain items cannot be carried along with them.

Data Mining:

Customer's profile update activity corresponding to the 'Country code change in the mobile number field' or any 'change in the zip code of the Address' field can be used as a trigger for this recommendation. And the products that are going to be recommended will be learned from the purchase history within 3 weeks from the move-in date of the customers in the past. The variables that need to be considered here are the Customer ID, Customer name, Gender, Country code of the Phone number, Zip code, Product purchased, Days from move-in date. The following database table shows a sample table with all these variables:

Training Set								
Customer ID	Name	Gender	Country code	Phone number	Address-Street	Zip code	Purchase History	Days from move-in
12345	John	M	+1	(813)-666-7343	Shiloh woods	33615	Sleeping bags, Mixers	2 Wks
			+61	453-695-482	Melbourne	3000		
14253	Pandur	M	+91	9392564098	Hyderabad	500084	Jackets	3 Wks
			+1	(814)-345-5447	Tampa	33613		
15685	Monica	F	+1	(907)-244-6437	New York	12404	Laptop	1 Wk
			+1	(907)-244-6437	Chicago	11453		
:	:	:	:	:	:	:	:	:
Testing Set								
16845	Jessica	F	+61	452-678-983	Sydney	2000	???	
			+1	(709)-356-5493	Atlanta	30301		
17543	Jacob	M	+1	(452)-978-9831	Utah	84001	???	
			+1	(452)-978-9831	Atlanta	30311		

Evaluation of the model:

In the above sample table, we could see that all the records are grouped into the Training and Testing sets. The split ratio will be chosen by us. In the training set, certain percentage of records of the historical data are used to train the model using various algorithms and the system is made to learn patterns. And then the remaining records are used for testing the accuracy of the model by making the model to predict the products and comparing that to the actual ones. Once the best possible algorithm is selected and the model is built the customers transferring to the new place are identified using the aforementioned profile update activities and our new category will be recommended to them on their home page of the website. Also, offers or discounts will be given on these items when bought together, thereby, preventing the customers to buy the small or easily available items in the local stores as soon as they move in.

Cost for the model:

As Amazon already has a complex system in place and are having huge volumes of user data available with them, adding this small model shouldn't take much of the efforts and cost much to the organization.

Integrating the data mining model:

The important thing to be noted here is that the proposed recommendation model is not a replacement to the existing Amazon's recommendation models. Rather, it is a new model that

can be built and applied to specific target customers (those who shift to a new place in our case), in addition to the regular recommendations that Amazon makes. This can be made to appear below the search based recommendation space on the home page, with the title 'Suggestions based on your recent activity' or can be sent as a message or as an email sending a link to the customers and asking them to have a look at the suggested products that they may like.