Article can be accessed online at http://www.publishingindia.com

Restaurant Recommendation System

Surajit Ghosh Dastidar*

Abstract

In the present paper a restaurant recommendation system has been developed that a recommends a list of restaurants to the user based on his preference criteria. There are two kinds of data files that have been used: restaurant master and customer master. Restaurant master consists of restaurant specific data and customer master consists of customer specific data. We have used decision tree algorithm to classify the customers into high, medium and low budget buckets based on customer demographics and purchase behaviour variables. Similarly, restaurants are also classified based on price category. The rules given by the decision tree algorithm are fed into a dashboard designed using MS Excel. The user can use this dashboard to get a list of restaurants based on his individual preference. The restaurant list is sorted based on users location details with the closest restaurant coming at the top of the list.

Keywords: Restaurant, Recommendation System, Decision Tree

Introduction

The concept of automated recommender systems was introduced in the 1990's and since then have revolutionised e-commerce by providing personalised recommendations and predictions over a variety of large and complex product offerings. A recommender system is any system that produces individualised recommendations as output or has the effect of guiding the user in a personalised way to interesting or useful objects in a large space of possible solutions (Burke, 2002). Recommendations made by such systems can help users navigate through large information about product descriptions, news articles or other items. Apart from user product preferences and taste, location based recommender system also exists which provides recommendation based on location preferences.

Literature Review

Recommender systems are widely popular in e-commerce environment. For example, UrbanSpoon allows online users to see which restaurants have been visited frequently by friends and relatives. Users can also specify the type of cuisine and price range to select their preferred restaurants. Yelp provides location, user reviews to find their restaurant of choice. Rinner and Raubal (2004) designed Hotel Finder which recommends hotels based on users location, spatio-temporal constraints, and specification. Ringo music recommender system allows users to express their musical preferences by rating various artists and albums and also suggests based on similar preferences of other users. FourSquare recommends restaurants based on user-specified criteria such as rankings and ratings. Maes, Guttmann, and Moukas (1999) provides a recent survey of recommendation systems.

Recommender systems can broadly be categorised into the following:

Collaborative or Social Filtering Recommender Systems

These systems aggregate data about customers purchasing habits or preferences, and make recommendations to other users whose profile matches with the past existing users. They basically use the concept of profiling. Collaborative systems works by building predictive models in order to predict the estimated of how much the user will like the set of items listed.

Content based Recommender Systems

These systems recommend items based on user item description and user profile (demographic, purchase). Content-based recommendation systems are used in

* IMT Hyderabad, Hyderabad, Telangana, India. Email: sghoshdastidar@gmail.com

various domains ranging from recommending web pages, news articles, restaurants, television programs, and items for sale. It maintains the user profile in the database and updates the same on regular basis. These systems use supervised learning concept in order to predict the most likely results of a query given by the consumer. Content-based recommendation system analyse item description in order to identify items that are in line with the interest of the user.

Knowledge based Recommender Systems

This kind of systems uses knowledge about the users and products to pursue a knowledge-based approach to generate recommendations for a particular user and product.

Objective

The objective of this paper is to design and implement a localised, personalised, and content-based recommender system for restaurants. It should be easy to understand and implement.

Data Source

The data was obtained from University of California Irvine machine learning repository. There were nine data files which were categorised into restaurants, customer and customer ratings.

Data File for Restaurants

- Payment information
- Cuisine information
- Operating hours information
- Parking facility information
- Geographical and profile information

Data Files for Customers

- Customer cuisine
- Customer payment
- Customer profile

Data Files for Customer Ratings

Ratings given by the customers to each restaurant based on three parameters namely ratings, food and service.

Two master files namely restaurant master and customer master were created from the above mentioned eight files. Restaurant master file contained all the information regarding a particular restaurant. There were a total of 95 restaurants in the file. Customer master file contained all the information regarding a particular customer. There file had information of 135 customers. Table 1 gives the details of all the variables for each of the files.

Table 1: Details of Master files

Customer Master	*Restaurant Master*
userID	placeID
Latitude	Rcusine
Longitude	Rpayment
Smoker	Parking
drink_level	Weekday time
dress_preference	Sat
Ambience	Sun
Transport	Latitude
marital_status	Longitude
Hijos	Name
birth_year	Alcohol
Interest	smoking_area
Personality	dress_code
Religion	Accessibility
Activity	Price
Color	Rambience
Weight	Franchise
Budget	Area
Height	other_services
Cuisine	
Payment	

Methodology

Literature review suggests that decision tree is the most preferred technique for content-based recommendation system. Hence, decision tree was chosen for categorising the customers and the restaurants. The decision tree algorithm was implemented using SAS Enterprise Miner 4.1

Results

Descriptive Analysis

It has been observed that 69% of the customers had a medium budget while only 4% had a high budget (Fig. 1). Mexican cuisine is offered by almost 29% all the restaurants (Fig. 2). 91% of the customers preferred cash transactions (Fig. 3). Around 45% of the restaurants did not have any parking facility (Fig. 4) while 33% offered alcohol services (Fig. 5). More than 60% of restaurants had smoking facilities (Fig. 6). The dress code for 90% of the restaurants was informal (Fig. 7).

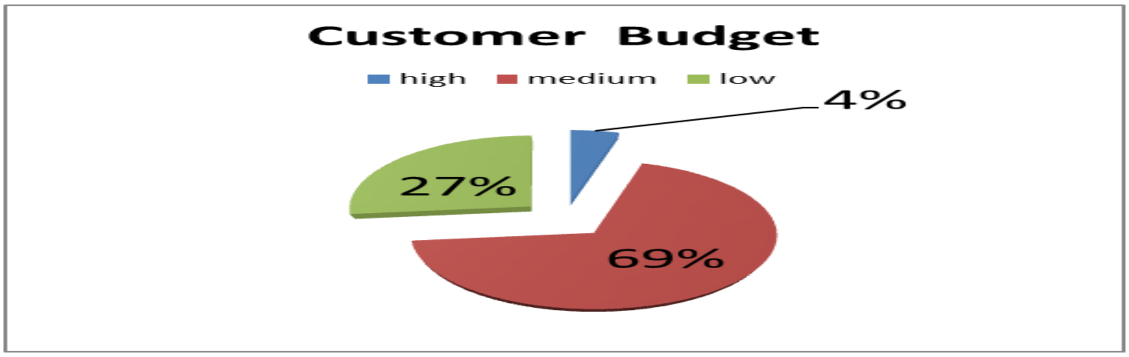

Fig. 1: Customer Budget

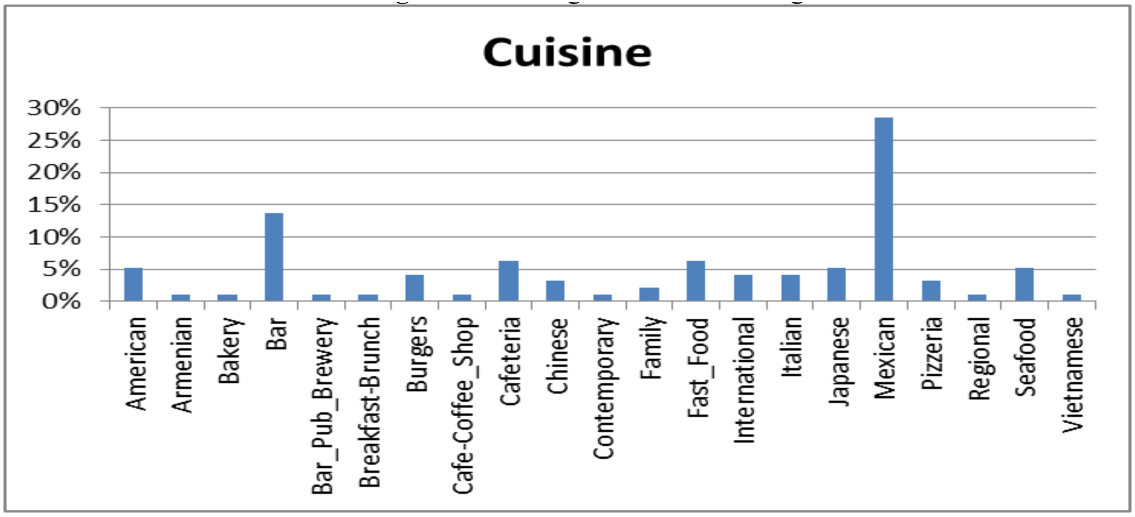

Fig. 2: Cuisine

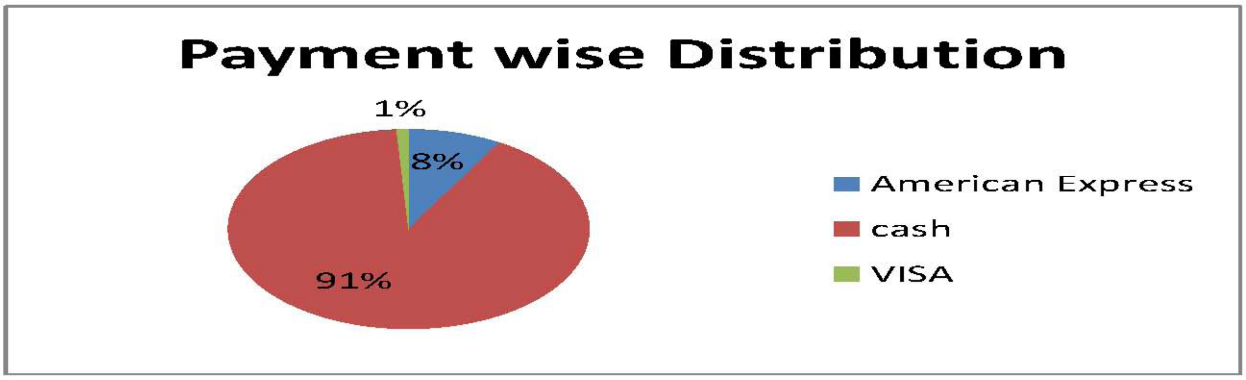

Fig. 3: Payment wise Distribution

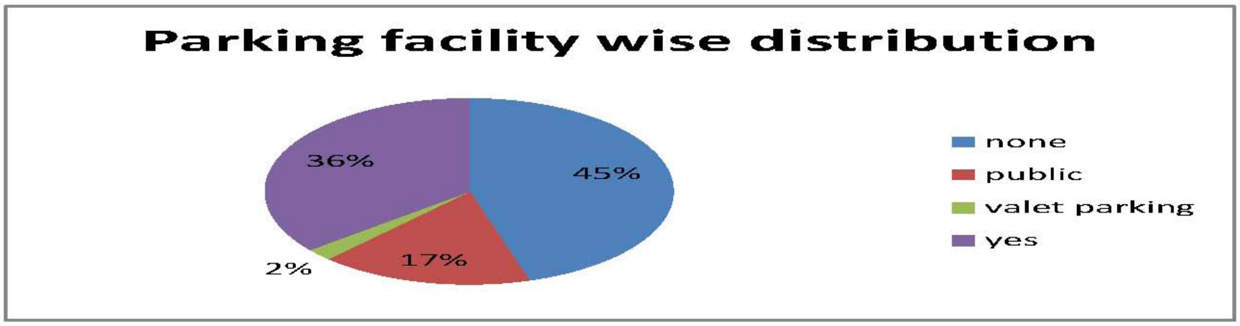

Fig. 4: Parking Facility wise Distribution

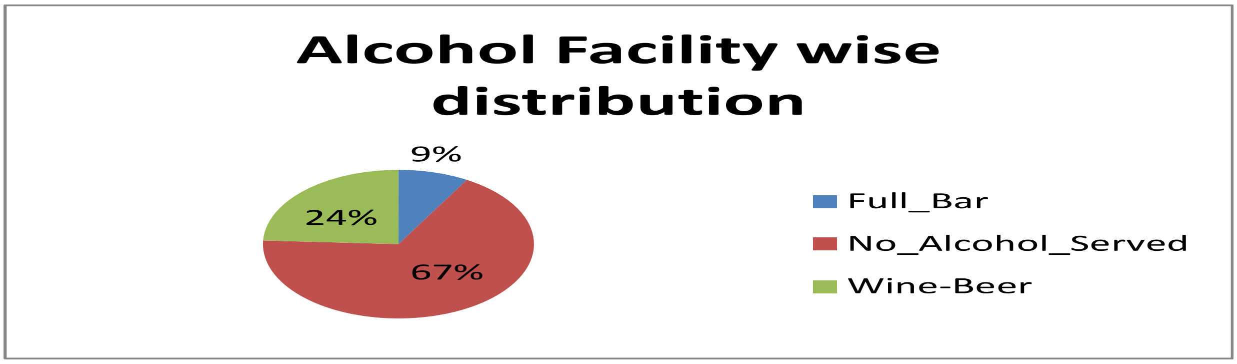

Fig. 5: Alcohol wise Distribution

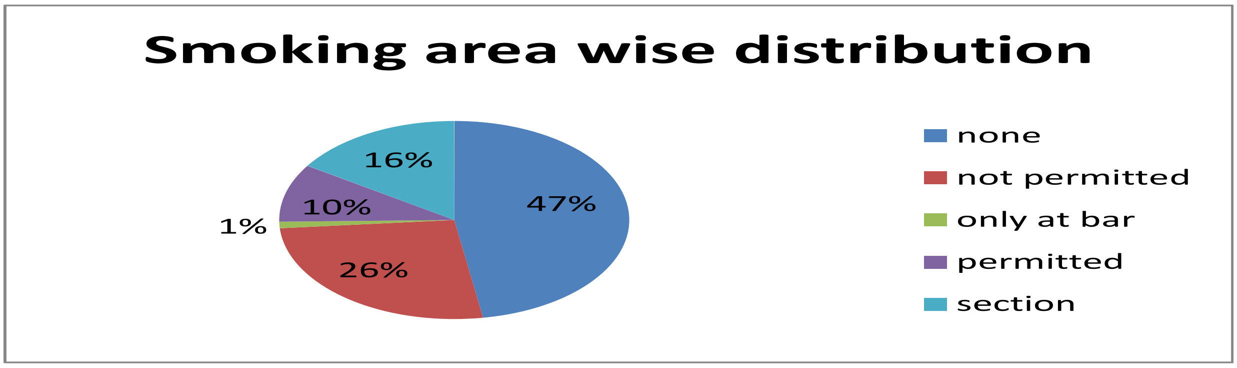

Fig. 6: Smoking area wise Distribution

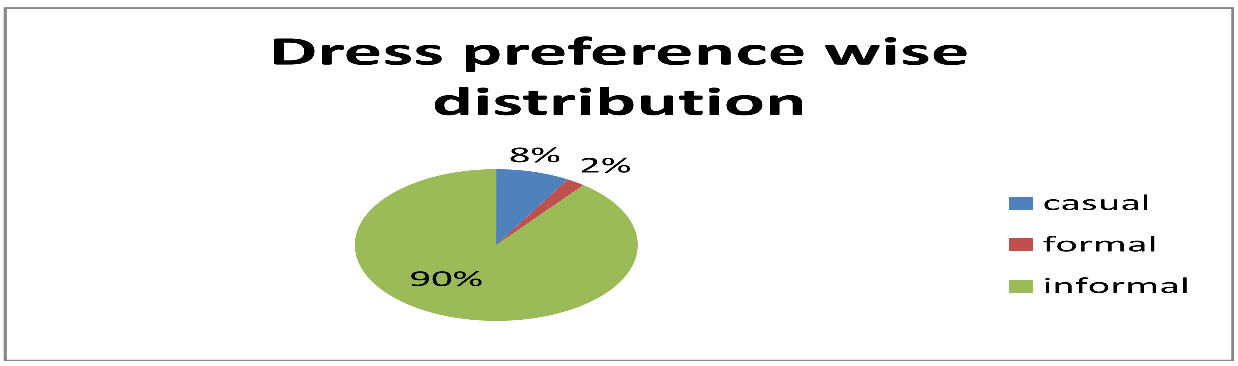

Fig. 7: Dress Preference wise Distribution

We also clustered the restaurants into three clusters based on the average user ratings for each of the restaurants. The following table summarises the results.

Table 2: Ratings of Clusters of Hotels

Avg User Rating	*Cluster No.*	*High*	*Low*	*Medium*	*Grand Total*
2.49	1	7%	64%	29%	28
4.75	2	20%	20%	60%	25
3.51	3	17%	26%	57%	42

Cluster 1 had 64% restaurants in the low budget category. Interestingly, the average user ratings for this cluster was also found out to be 2.49. Cluster 2 had 80% of the restaurants in either high or medium priced category and had the maximum average user rating of 4.75. In cluster 3 the average user rating was 3.51 which constituted 74% of the restaurants belonging to medium or high priced restaurants. Hence, it can be concluded the majority of the low priced restaurants belonged to cluster 1 which also had a low user ratings while cluster 2 had the highest average user rating of 4.75. From this analysis we can conclude that low budget customer tend to go to low priced restaurants while high budget customer tend to go to high priced or medium priced restaurants. The first preference for medium budget customers is medium priced restaurants followed by low priced restaurants. In this study we have assumed a one to one relation between customer budget and restaurant category as stated earlier.

All the restaurants were classified into low, medium and high categories based on price. The restaurants were also classified into three categories "best", "ok", and "not ok" based on services offered, food quality, and review. It was observed that the number of "not ok" restaurants was maximum in the low price category compared to medium and high priced categories. It was interesting to note that none of the high priced category restaurants belonged to the "not ok" category.

The decision tree was obtained for the customers as shown in Fig. 8. The misclassification rate was found to be 0.20. Only those rules were considered for developing the dashboard where the probability of categorising the customers in any of the low, medium or high budget categories were greater than 70%

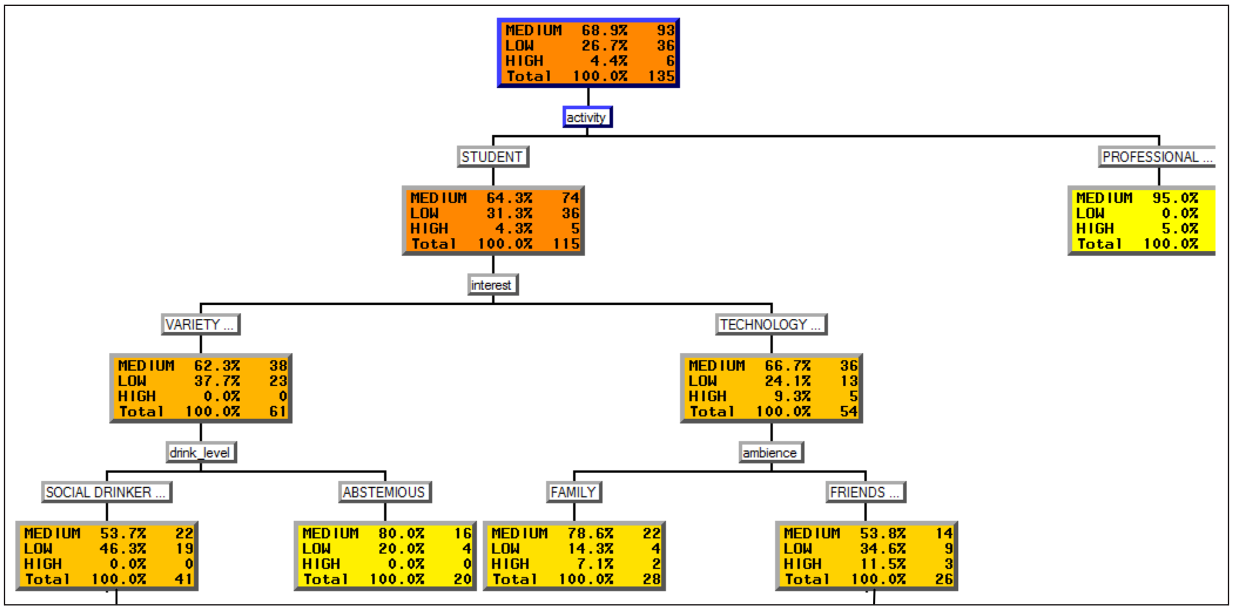

Fig. 8: Decision Tree for Customers

The decision tree was obtained for the restaurants as shown in Fig. 9. The misclassification rate of the following tree was found to be 0.32. We have considered only those rules where the probability of categorising the restaurants into low, medium or high priced restaurant is more than 0.70.

The decision rules will help the existing restaurant owners to identify what differentiates high priced restaurants from medium or low priced ones. Accordingly, the restaurant owner can add relevant features and functionalities to his existing restaurants in order to attract high budget customers.

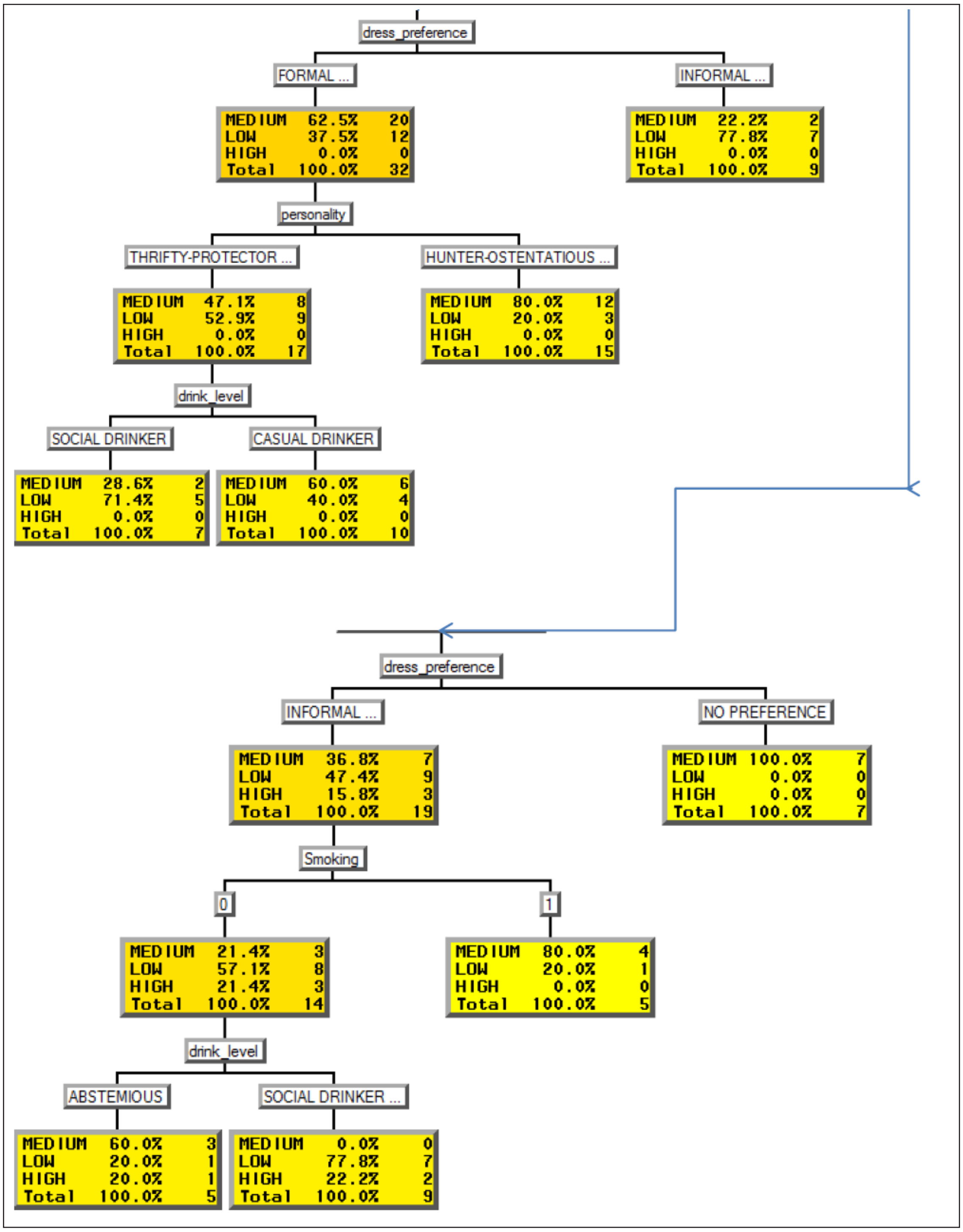

Fig. 9: Decision Tree for Restaurants

We have designed a dashboard in excel which is primarily meant for the customers. The customers can enter his personal preferences for a restaurant based on choices like smoking, interests, drink level, dress preference, ambience, personality and activity. The underlying decision tree rules will categorise the customer as high, medium or low budget customer and display a list of all the restaurants belonging to a particular category. If the customer further enters the location details of him in terms of latitude and longitude, the existing list of restaurants is further sorted in ascending order with the nearest restaurant coming at the top of the list.

Table 3 shows the input format for entering customer preferences. The customer preferences have to be entered in binary format, 1 meaning TRUE and 0 meaning FALSE.

Table 3: Input Format for Entering Customer Preferences

Customer specifications Input				
Smoking	FALSE	TRUE		
	0	1		
Interest	variety	retro	eco-friendly	technology
	0	1	1	1
drink level	abstemious	casual drinker	social drinker	
	0	0	1	
dress preference	formal	informal	no-preferecne	elegant
	0	1	0	0
ambience	family	friends	solitary	
	1	1	0	
personality	conformist	hard-worker	hunter-ostentatious	thrifty-protector
	0	1	0	0
Activity	student	proffesional	unemployed	workingclass
	0	1		0

	User Location details
Latitude	180
Longitude	102

Table 4 shows the output which is a list of restaurants along with restaurant details and distance from customer's existing location as specified by the latitude and longitude.

Table 4: List of Restaurants along with Other Details

Restaurant Recommendation system							
Restaurant Recommendation system	*Distance*	*Cuisine*	*Weekday*	*Saturday*	*Sunday*	*Parking*	*Area*
Name	*Distance*	*Cuisine*	*Weekday*	*Saturday*	*Sunday*	*Parking*	*Area*
Restaurant Bar Coty y Pablo	1,941.04	Bar	16-24	16-24	16-24	none	closed
El cotorreo	1,943.37	Family	0-8	0-8	0-8	none	open
Cafeteria cenidet	1,944.12	Cafeteria	0-8	0-8	0-8	public	closed
Log Yin	1,946.28	Mexican	16-24	16-24	16-24	yes	closed
Subway	1,946.80	Fast_Food	16-24	16-24	16-24	public	closed
McDonalds Centro	1,946.90	American	16-24	16-24	16-24	none	closed
TACOS CORRECAMINOS	2,169.71	Mexican	16-24	16-24	16-24	none	closed
TACOS EL GUERO	2,170.56	Mexican	16-24	16-24	16-24	none	closed
tacos de la estacion	2,170.57	Mexican	0-8	0-8	0-8	none	open
little pizza Emilio Portes Gil	2,171.78	Armenian	16-24	16-24	16-24	none	closed
palomo tec	2,171.84	Mexican	16-24	16-24	16-24	none	closed
tacos de barbacoa enfrente del Tec	2,171.86	Mexican	8-16	8-16	8-16	public	open
Carreton de Flautas y Migadas	2,171.90	Mexican	0-8	0-8	0-8	none	open
puesto de gorditas	2,171.92	Regional	0-8	0-8	0-8	public	open
tacos abi	2,171.93	Mexican	16-24	16-24	16-24	none	closed
Hamburguesas La perica	2,172.16	Mexican	16-24	16-24	16-24	public	open

		Distance(miles)
Nearest Restaurant	Restaurant Las Mananitas	1,947

Conclusion

The paper discusses about the design and development of a localised, personalised and content-based recommendation system for restaurants which is easy to understand and implement using MS Excel. However, the location finder is limited in its functionalities since the location of the customer has to be entered manually. In an online or mobile environment the system can be designed to capture the location details from the internet address or the in-built GPS embedded in most of the mobile phones. In addition, the accuracy of the decision tree is limited to the small amount of customer data. Further testing needs to be done with larger database of customers and restaurants.

References

Burke, R. (2002). Interactive critiquing for catalog navigation in e-commerce. *Artificial Intelligence Review*, 18(3-4), 245-267.

Maes, P., Guttmann, R. H., & Moukas, A. G. (1999). Agents that buy and sell, *Communications of the ACM*, 42(3), 81-91.

Rinner, C., & Raubal, M. (2004). Personalised multi-criteria decision based strategies based on location based decision support. *Journal of Location Geographic Information Sciences*, 10, 149-56.

Web References

Foursquare.com

www.urbanspoon.com