Tele-market Modelling of Fuzzy Consumer Behaviour

R. Basha and J. Ameen

Abstract—Just like any other aspect of human life on our planet, the advancements in internet technology over the past decades have greatly influenced the way we live, changing the socio-political structure of the nations as it can clearly be seen in the Middle East and northern African countries. In specific terms, they have influenced our behaviour in the way that we conduct marketing extending it beyond traditional boundaries both from business actions and consumer's behaviour. With ease of access and the vast amount of information that can be observed on the internet, consumers are lost for choice. These have increased the challenges that both consumers and business people face manifold. The survival conditions for both sides require more information in order to build new models that can respond to these challenges and help the decision making process in line with these changes and making them more efficient.

This paper attempts to assess the fuzzy actions of buying behaviour from a multinational viewpoint using data that have been collected on consumer risk assessment while attempting to buy a product on the internet. The survey covered a sample of 270male and female participants of different nationality with different levels of income and education.

In a hierarchical modelling attempt, Logistic regression is used to identify factors that are significant in the action of purchasing leading to a sensitivity of 0.94 and specificity of 0.81. The resulting significant components from the first stage model are used as inputs in the application of AnswerTreeand more specifically, Classification Regression Tree (CRT) model to formulate what-if decision scenarios to help decision makers improve their targeting and selling options. The sensitivity and specificity of the latter approach were 0.93 and 0.83 respectively and led to the establishment of a set of decision rules each with their probability of success ranging from 0.67 to 1. None of the factors of gender, geographic location or income were significant in the process.

Keywords—Tele-marketing, Customer behaviour, Logistic Regression, Decision Tree

Manuscript submitted May X, 2011:

R Basha is Associate Professor at the College of Computer Information Technology, American University in the Emirates (AUE)
Dubai International Academic City-Dubai UAE (email: rawshan.basha@aue.ae)

J Ameen is Professor of Statistics at the Department of Mathematics and Computing, University of Glamorgan, Pontypridd, South Wales, UK, CF37 1DL (http//staff.glam.ac.uk/users/3851-jrmameen) & Senior Advisor, Kurdistan Regional Government, Erbil, Iraq. (Phone: +9647504288482; email: jrmameen@glam.ac.uk).

I. INTRODUCTION

ITis well recognized that recent advancements in communication technology have influenced every aspect of societies across the globe. It is no longer viable for businesses to restrict themselves to their classical market boundaries within districts, regions, countries and even continents. This is part of the issue of globalization that has been addressed in the literature as a means of survival not only for businesses but also the socio-political aspects of a society (Orozalieva 2010).

This is mainly due to the fact that customers are able to communicate with information and goods from anywhere on this planet provided that relevant facilities are available at the two ends for safe and reliable exchange of such information and goods. Techniques in the display of goods and products in local stores, while remain important, at a higher level, these have now been translated to ease of access, well designed and attractive websites from which factors like security, reputation, trust, privacy and pricing can be of influence in consumers buying decision. The causality of either of these factors in the final consumer decisions is never a clear cut and hence a loose probabilistic nature for ease of effective decision making would be desirable. It is also important to note that while a number of factors have been important in attracting consumers and affecting their choices during the beginning of the trend in adopting internet facilities for shopping, with the advent of time, these factors have been subjected to change and neither the consumer behaviour nor the influence of these factors on the former are expected to remain unchanged.

II. BACKGROUND

The challenges that advancements in information technology have brought about, are immense since its start up about two decades ago. Perhaps the most significant impact it has made, has been in the area of communication and its allied areas (Turban et al., 2000) which in turn has changed the social and political maps on our planet for good. This is evident in the way we see the map of the Middle East and Northern African countries are beginning to change irrespective of the resistance to them. These advancements have caused regions and countries boundaries to loosen their social and political boundaries and more specifically has greatly influenced the marketing process in that it has changed the marketing frames and expanded them beyond their countries of origin and businesses have established platforms from which they display their products worldwide (Liu and You 2003). In contrast, consumers, irrespective of their geographic location, have now the opportunity to shop on the internet and this has made it more challenging for analysts and modellers to develop efficient marketing strategies to account for these changes and accommodate them in their newly revised models as most of the earlier developed classical models redundant.

Among the modelling tools that have commonly been used in the literature, are regression based models(Ferraro et al. 2010, Yen et al. 1999, Tanaka and Watada 1998, Che-Chem et al. 2009, Xue-Cheng et al. 2008 and Tsang et al. 2011) and logistic regression models with multiple explanatory variables (Demirtas et al. 2009, Hosmer and Lemeshow 2000, Sarlija N et al. 2006).In a research study presented in one of the wseas conferences on applied mathematics, Yang and Chiu (2006) have adopted both the logistic regression and decision trees to identify factors that minimize customer attrition.

While a careful use of these models is necessary in identifying components that would be expected to have significant influences on the main parameter of interest, despite problems of over parameterization, their results stop short of being useful for decision makers at customer level. This is mainly the case due to the nature of the modelling approach based on a high level of the aggregated data for the sake of obtaining reliability and generalizability on one hand and the fuzzy nature of consumer demographics and behaviour on the other (Akhter et al. 2005). It is therefore of advantage to adopt any tools and models that could be seen to assist in getting the results of the modelling approach closer to decision makers wishes for taking actions at field level filling the gap that already exist in this process.

The latter problem is therefore addressed using a second stage modelling approach known as "AnswerTree" and in particular the Classification and Regression Tree (CRT) and Chi-squared Interaction Detection (CHAID) methods (Miloslova K & JIRI K 2008, Chin-Sheng, et al. 2008, Lemon et al. 2003, Mei-Ping and Wei-Ya 2010) which is one of the components that has been made available within the Statistical Packages for Social Sciences (SPSS). This allows for the creation of tree based classification models identifying the best 'in some pre-specified sense' possible scenarios through partitioning of the explanatory variables using Chi-Squared statistical technique to improve the predictive power of the explained variable.

This research is therefore based on adopting a hierarchical modelling approach through the application of binary logistic regression models to identify factors that are significant in guiding consumer behaviour at an aggregate level using data that have been collected as a result of a survey on 270 participants of different nationality, gender, income and levels of education in their attempt to buy products through either of the three websites: Lastminute.con, Amazon.com and Uaemall.com that have been made available to us by Akhter et al. (2005) with thanks. This step is then followed by the application of AnswerTree to produce decision rules each with their probability of success which is expected to fill the gap between statistical modelling outputs and outputs reflecting the fuzzy nature of consumers in their buying behaviour and possible actions that need to be taken at market levels and as part of marketing policy strategies. The importance and efficiency of multilevel modelling strategy in combining different modelling procedures to improve the presentation of consumer buying behaviour have been acknowledged by many researchers (Bae and Kim 2010). Multilevel modelling approaches have been used in many other areas. Pereira-Moliner et al. (2011) have used this approach to establish the relative importance of intergroup and intragroup performance differences based on the use of hierarchical linear models.

III. DATA CONSIDERATIONS

A web based survey had been conducted by Akhter et al. (2005), in which data were collected to assess human reasoning on electronic commerce transactions. This study did not attempt to learn about the buying behaviour of consumers on the internet although enough information was collected from 270 participants who were enrolled on Information Systems and Computer Science classes. Participants formed a mix of gender (146 for male & 124 for female), nationalities (93 Middle Eastern, 76 North Americans, 56 Europeans, 35 Africans and 10 Australian and South Americans), different levels of education with the majority (86%) holding BSc, MSc or PhD degrees and their annual income was fairly distributed between any amount "less than $25k" to any amount more than "$100k" and they were given credit and asked to go through the whole process of buying what they wish to buy, from the designated websites: Amazon.com, Lastminute.com and Uaemall.com.

In addition to the demographic information on each of the participants, information was recorded on:

1. Time each participant remained on the internet.
2. Number of times that transactions were cancelled.

Furthermore, questions were asked if:

1. The importance of the company having each of the website having a return policy, cancellation without penalty, privacy, discount, privacy, ease of use, reputation, the way the icons located and presented as the participants perceive.
2. They have purchased anything on the internet before.
3. Security features could be a factor in their buying behaviour.
4. They are happy to provide credit card numbers via non secure web pages to an on0line company.

The products that were available for consumers to buy and actually bought through the implementation of this exercise, were categorized into the following each with their percentage bought by the participants: CD Records (0%), Clothes (12.6%), Computers or Accessories (12.2%), Travel (20%), Books (31.5%) and a 23.7% of none of those.

IV. METHODS

This research has used results of a web-based survey after directing subjects who were enrolled on information systems and computer science classes to examine the buying process from a number of well-known e-sales websites. Buyers were left free to select the product they intend to buy and in addition to demographic information, their moves and perceptions were recorded from the start to the point of ordering the product.

Data were recorded in SPSS 15 and binary logistic regression was used to identify factors that are significant in the buying process. Just like multiple linear regression techniques, this approach is based on expressing the explained variable (but in natural logarithmic terms) as a linear function of the expected explanatory variables. The base of the explained variable is meant to have a binary format as is the case with buying/not buying actions (the explained variable, $= \frac{p}{1-p}$; where p being the probability of buying a product). This is then expressed as:

$$y = e^{b_0+b_1x_1+b_2x_2+\cdots+b_nx_n} + Error$$

Where $x_i\,; i = 1,2,\ldots n$ represent the explaining variables with their coefficients estimated from the collected data as listed in Table 1 below. Furthermore, it can be seen that the coefficients represent elasticities (That is, every b_i percentage of change in x_i would be expected to lead to one percentage of change in y) for their corresponding explaining variables as the following formula can demonstrate:

$$\frac{1}{y}\frac{\partial y}{\partial x_i} = b_i; i = 1,2,\ldots n$$

Furthermore, in the case of dichotomous variables, as expressed in table 1, the EXP(b_i) values, represent the odds ratios for their corresponding explanatory variables.

This modelling step was taken to identify significant factors that are influencing the buying behaviour each with the strength of their contributing to that process in order to pave the way for the application of the second stage modelling approach to get closer to the decision making at base level.

Following the identification of significant factors influencing the buying behaviour using the binary logistic regression approach as described above, AnswerTree and specifically the Classification and Regression Tree (CRT) was used to develop a fuzzy model to assist targeting sales through what-if buying decision rule scenarios each with their probability of success.

It is important to establish that this research has focused on identifying predictor variables that are likely to contribute or contributing to the action of buying a product on the internet. Although these variables could also be causal, they may not be so. In more general terms, not all predictor variables are causal but all causal variables are predictors as well. The establishment of causality could be addressed under a different study and requires more analytical foundation and time and goes beyond the current brief (Granger 1969 and Ameen and Naji 2000).

V. RESULTS

A. Binary Logistic Regression

In an attempt to explain the buying behaviour of consumers at an international level, Binary Logistic Regression models were adopted expressing the buying action as the explained variable with the expected explaining variables in three groups of demographic (including gender, age, level of education, annual income, country of origin), personal attitude (including the importance each participant give to specific website policies) and issues related to website appearances all as indicated in section III above.

Following these principles and using the collected data, the resulting binary logistic regression model identified the variables that are significant in predicting the buying behaviour as indicated in table 1 below:

Table 1 Logistic Regression Results

Model Factors	B	S.E.	Wald	df	Sig.	Exp (B)	L	U
Time On Internet	-0.9	0.3	9.7	1	0	0.39	0.22	0.71
Discount	-0.6	0.2	6.8	1	0.01	0.56	0.36	0.87
Online safe	0.5	0.2	5.3	1	0.02	1.71	1.08	2.69
Security Features	-0.9	0.4	6.4	1	0.01	0.39	0.19	0.81
Purchase Online	1.1	0.2	39.8	1	0	2.89	2.08	4.01
Tran terminated	0.4	0.1	8.9	1	0	1.47	1.14	1.9
Important Factor	0.2	0.1	4.3	1	0.04	1.26	1.01	1.57
Constant	-4.3	1.4	9.9	1	0	0.01		

L: Lower 95% Confidence Level; U: Upper 95% Confidence Level.

As expressed in Table 1 above, the most significant factor in determining the buying action turned out to be the type of product that has been described for sale by an odds ratio of near three to one (2.89).

In terms of probabilities, this result means that the estimated probability of buying the product based on the nature of the product that has been put for sale would easily be calculated from the formula:

$$\frac{p}{1-p} = 2.89$$

From which $p = \frac{2.89}{3.89} = 0.743.$

Similarly, this is followed by the buyers feel factor for safety with an odds ratio of 1.71 (p=0.631) followed by “Transaction Terminated” with an odds ratio of 1.47 (p=0.595).

The factors that correspond to odds ratios of less than one, (Exp(B) in Table 1), are indicated to be inversely related to the buying behaviour. For instance, the more the participants remain on the internet investigating to take a buying action, the less likely it is that (s)he will make a purchase. The more conserved the participant is about issued of discount and security features, again, the less likely that they make a purchase.

In terms of probabilities, this means that the probability of the occurrence of the event would be less than 0.5, which is less than the probability of its non-occurrence as they both have to complement each other. On this basis, each of the factors of “Time on the Internet”, “Discount” and “Security Features” is of this nature.

Model validity is expressed in Table 2 as an output of the SPSS model validation. Although both measures of Cox and Snell and Negelkerke R squares are expressed as common measures of the strength of association as reported in the SPSS output, Negelkerke R Square measure is more common and considered as a better measure of the strength of association. In our modelling case, an indication of the capturing of near 70% of the variability in the data as expressed by the Nagelkerke R^2 value is a good indication of model validity (Table 2).

Table 2 Model Summary

-2 Log Likelihood	Cox & Snell R Square	Negelkerke R Square
138.624	0.455	0.675

Model predictive power is explained by Hosmer and Lemeshow test (table 3) aggregating observations in to groups of similar cases. The model has been able to predict both purchases and non-purchases in all cases in a highly satisfactory manner.

Table 3 Contingency table for Hosmer and Lemeshow Test

Purchase On Internet = yes		Purchase On Internet = no		Total
Observed	Expected	Observed	Expected	
27	26.930	0	.070	27
26	26.745	1	.255	27
26	26.527	1	.473	27
27	26.190	0	.810	27
27	25.655	0	1.345	27
25	25.560	3	2.440	28
23	22.702	4	4.298	27
15	15.540	12	11.460	27
5	5.897	22	21.103	27
2	1.253	24	24.747	26

This model validity is further emphasized by the results obtained from the predictive model using the well-known measures of sensitivity and specificity which provide a further set of simultaneous measures of model reliability in predicting the true cases correctly as being true and the false cases as false. That is correctly predicting the true buying action of a buyer and also the specificity measure that is indicating the reliability of the model in predicting that the buyer will not proceed to buy. The probabilities of 0.94 and 0.81 are highly acceptable for sensitivity and specificity respectively (See table 4). These are in contrast with the commonly used and well known type-1 and type-2 errors in hypotheses testing as a main component of statistical decision theory.

Table 4 Logistic Regression Predictive Quality

Model predictions		Predicted		
		yes	No	Sensitivity
TRUE	Yes	191	12	0.94
	No	17	50	
	Specificity		0.81	

These models are very useful in building and explaining the overall buying environment and in the identification of significant factors that influence the buying behaviour. However, they remain short of having recipes reflecting consumer behaviour to help marketing decision makers set plans to better target and improve their marketing and selling strategies.

B. Decision Tree

SPSS AnswerTree can be viewed as a powerful data mining tool that can help identify segments of large datasets that can lead to better predictions of a certain pre-specified factor. One of the methods for achieving such segmentation is called the Classification and Regression Tree (CRT) that is based on the classification of explaining variables to ‘best’ predict the explained (target)variable in the sense that such partitions are as homogeneous as possible with respect to the explained variable (Fujii1992, Hilb 2005). This will lead to the best partition based on the pre-set and accepted statistical significance level creating a tree-like partition from all the factors of the dataset that turn to be significant in predicting the explained variable.

In this paper the CRT procedure is applied within a hierarchical modelling strategy taking the identified factors that are expected to significantly influence the buying behaviour as a result of the application of the binary logistic regression model. In that sense, the predicted variable has been set as the internet buying indicator but restricted the explained variables only to those identified as significant through the application of Logistic Regression with the ‘Gini’ impunity measure making sure that the splits maximize the homogeneity measure of the child nodes with respect to the value of the target variable; maximum tree depth of five;

minimum cases in parent node to be 15 and minimum cases to be 10 in child nodes as initial parameter setups.

A summary of the model specifications are given below in table 5.

Table 5 CRT Model Summary

Specifications	**Growing Method**	**CRT**
	Dependent Variable	PurchaseOnInternet
	Independent Variables	TimeOnInternet, Discount, Onlinesafe, SecurityFeatures, PurchaseOnline, Tranterminate ImportantFactor
	Validation	None
	Maximum Tree Depth	5
	Minimum Cases in Parent Node	15
	Minimum Cases in Child Node	10
Results	Independent Variables Included	PurchaseOnline, Onlinesafe, Tranterminated, Discount, TimeOnInternet, ImportantFacto SecurityFeatures
	Number of Nodes	17
	Number of Terminal Nodes	9
	Depth	5

Within the main model specifications in this case, the depth of branching the developed tree was subjectively set so that they will not exceed five and parent nodes were allowed to have a minimum of 15 cases and child nodes to have a minimum of 10 cases.

The resulting decision tree further identified the factors that are significant in determining the buying behaviour together with the order and degree of their significance as presented in the histogram below.

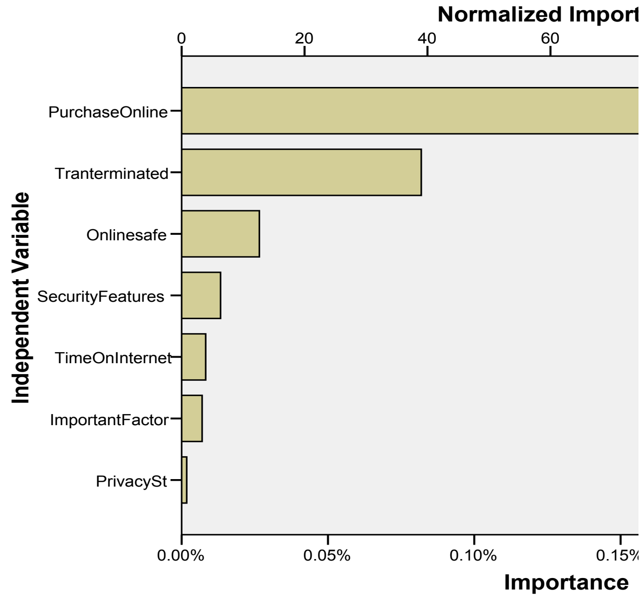

Figure 1 Ranked significant Regression Factors

It is important to note that the factors that have further been identified to be determining the buying behaviour are in the first place the type of the product that has been put for sale or the brand (Purchase on Line). This relates mainly and determined by the seller while the rest of the factors are all consumer or buyer related. For example, each of the factors: Transaction Terminated, Online Safe, Security Features, Time on the Internet, Important Factor which relates to consumers identifying the importance of ease of use, pricing, convenience, wide selection, trust, customer services, reputation, security, privacy and the option of none of these, as expressing how buyers feel confident about using the electronic systems for their purchasing activity.

The predictive quality of the resulting decision tree is expressed through the sensitivity and specificity measures as shown in Table 6. The sensitivity measure indicates the percentage of true buyers identified by the model as true buyers and the specificity measure of 83% represents those true non-buyers identified correctly as non-buyers.

Table 6 Decision Tree Predictive Quality

Model predictions		Predicted		
		Yes	No	Sensitivity
TRUE	Yes	192	11	0.93
	No	14	53	
	Specificity		0.83	

Both percentages of 93% and 83% are highly acceptable in practice and their closeness with their counterparts in the application of the logistic regression is an indication of consistency of the findings and the quality of the developed model.

The results of both tables 4 and 6 can further be explored to obtain two more useful measures of Positive Predictive Value (PPV) and Negative Predictive Value (NPV). These are described as:

$$PPV = \frac{True\ Positive}{True\ Posotive + Fake\ Positive}$$

And

$$NPV = \frac{True\ Negative}{True\ Negative + Fake\ Negative}$$

On the basis of Table 6, PPV= 0.935 and NPV=0.859.

The decision tree below clarifies the way in which decision makers can use the results of the AnsweTree classification to make decisions. These possible decisions will be in the format of what-if scenarios expressed through each of the bullet points listed in the appendix and depicted in the decision tree as in Figure 2.

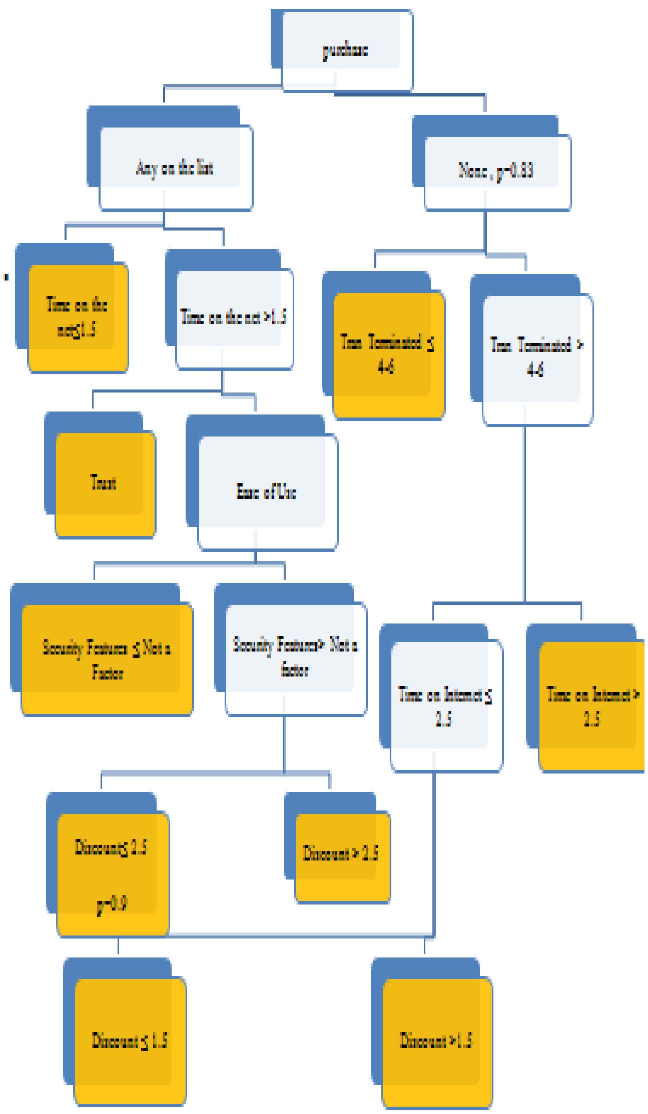

Figure 2 The Resulting CRT Decision Tree

V. Conclusion

The complexity of identifying grand rules on consumer behaviour within the current framework and the speed that tools and techniques of marketing expand and develop and the expansion in the ways that consumers access products and have the options of buying is beyond any doubt a big challenge for modellers and market analysts alike. While this paper is an attempt to address some of these issues, there is much left to be done and on a continuous basis in order for consumers as well as business managers and decision makers cope and respond to the dynamic changes that are taking place in the world of communication and information technology on one side and the changes in consumer behaviour as a result of access to information. The modelling results show that the main and the most important factor that has governed the process of buying within this group of participants is the type of product that has been put for sale followed by the buyers attitude reflected in the number of times that (s)he has terminated the process. All the other factors, though they show small degrees of significance, they disappear with time and improvements in advancements in internet technology in reassuring buyers.

Appendix

- P[((Purchase Online =2 OR Purchase Online =3 OR Purchase Online =4 OR Purchase Online =5) OR Purchase Online≠7 AND(Online Safe<3 OR Tran Terminated <3) AND (Ease of Use<4.5 AND Time On Internet<1.5))]= 0.72;
- P[((Purchase Online =2 OR Purchase Online =3 OR Purchase Online =4 OR Purchase Online =5) OR (Purchase Online≠7 AND Online Safe<3 AND Tran Terminated<3 AND Ease of Use < 4.5 AND Time On Internet>1.5))] =0.89;
- P[((Purchase Online =2 OR Purchase Online =3 OR Purchase Online =4 OR Purchase Online =5)OR (Purchase Online≠7 AND Online Safe<3 AND Tran Terminated<3 AND Ease of Use)<4.5 AND Time On Internet>1.5 AND Important Factor≠5 AND Security Features<1))]= 0.90;
- P[((Purchase Online =2 OR Purchase Online =3 OR Purchase Online =4 OR Purchase Online =5 OR Purchase Online≠7)AND (Online Safe<3 AND Tran Terminated<3 AND Ease of Use)<4.5 AND Time On Internet>1.5 AND Important Factor≠5 AND Security Features>1 AND Ease of Use)<2.5))]= 1.00;
- P[((Purchase Online =2 OR Purchase Online =3 OR Purchase Online =4 OR Purchase Online =5 OR Purchase Online≠7)AND (Online Safe<3 AND Tran Terminated<3 AND Ease of Use <4.5 AND Time On Internet>1.5 AND

Important Factor≠5 AND Security Features>1 AND Ease of Use)>2.5))] = 0.90;

- P[((Purchase Online =7 OR (Purchase Online≠2 AND Purchase Online≠3 AND Purchase Online ≠4 AND Purchase Online ≠ 5 AND Online Safe>3 AND Tran Terminated>3 AND Ease of Use>4.5))AND Tran Terminated<3 AND (Important Factor EQ 1 OR Important Factor =3 OR Important Factor≠2)AND (Important Factor≠4 AND Important Factor≠5 AND Important Factor≠10) AND Ease of Use > 4.5)] = 0.67;
- P[((Purchase Online =7 OR (Purchase Online≠2 AND Purchase Online≠3 AND Purchase Online≠4 AND Purchase Online≠5 AND Online Safe>3 AND Tran Terminated>3 AND Ease of Use >4.5 AND Tran Terminated>3)AND (Important Factor =2 OR Important Factor =4 OR Important Factor =5 OR Important Factor EQ 10) OR (Important Factor≠1 AND Important Factor≠3) AND (Ease of Use <4.5 AND Time On Internet<2.5 AND Important Factor ≠ 3 AND Important Factor ≠4 AND Ease of Use <1.5 AND Security Features<2) OR (Important Factor≠1 AND Important Factor≠2 AND Important Factor≠10))] = 1.00;
- P[(Purchase Online =7 OR (Purchase Online≠2 AND Purchase Online≠3 AND Purchase Online≠4 AND Purchase Online≠5 AND Online Safe>3 AND Tran Terminated>3 AND Ease of Use >4.5 AND Tran Terminated>3)AND (Important Factor =2 OR Important Factor =4 OR Important Factor =5 OR Important Factor = 10 OR Important Factor≠1)AND (Important Factor≠3 AND Ease of Use<4.5 AND Time On Internet<2.5 AND Important Factor≠3 AND Important Factor≠4 AND Ease of Use >1.5 AND Security Features>2 AND Important Factor = 1)OR (Important Factor =2 OR Important Factor = 10))] = 0.92;
- P[(Purchase Online =7 OR (Purchase Online≠2 AND Purchase Online≠3 AND Purchase Online≠4 AND Purchase Online≠5 AND Online Safe>3 AND Tran Terminated>3 AND Ease of Use >4.5 AND Tran Terminated>3) AND (Important Factor =2 OR Important Factor =4 OR Important Factor =5 OR Important Factor EQ 10 OR Important Factor≠1) AND (Important Factor≠3 AND Ease of Use <4.5 AND Time On Internet>2.5) AND (Important Factor =3 OR Important Factor =4))] = 0.82.

Acknowledgment

We would like to express our gratitude to Dr F Akhter [13] for providing the data that we have used in this paper.

.

References

[1] L.A. Zadeh (1987) Selected Papers on Fuzzy Sets and Applications. ed. R.R. Yager et al. John Wiley, New York.

[2] E. Turban, J. Lee, D. King and H. Chung, Electronic commerce: A managerial perspective, Prentice Hall (2000).

[3] J Liu and J You (2003) Smart shopper: An agent-based web-mining approach to Internet shopping. IEEE Transactions on Fuzzy Systems, 11(2), 226-237.

[4] M V Diepen and P H Franses (2005) Evaluating chi-squared automatic interaction detection. Information Systems, 31(8), 814-831

[5] M B Ferraro, R Coppi, G GRodriguez and A Colubi (2010) A linear regression model for imprecise response. International Journal of Approximate Reasoning, 51(7), 759-770.

[6] H Tanaka and J Watada (1998) Possibilistic linear systems and their applications to the linear regression model. Fuzzy Sets and System, 27, 275–289.

[7] K K Yen, S Ghoshray and G Roig (1999) A linear regression model using triangular fuzzy number coefficients. Fuzzy Sets and Systems, 106, 167–177.

[8] D W Hosmer and S Lemeshow (2000) *Applied Logistic Regression*, 2nd ed. Chichester, New York: Wiley.

[9] E A Demirtas, A S Anagun, G Koksal (2009) Determination of optimal product styles by ordinal logistic regression versus conjoint analysis for kitchen faucets. International Journal of Ergonomics, 39(5), 866-875.

[10] J R M Ameen and J A Naji (2000) Causal models for road accident fatalities in Yemen. Accident Analysis and Prevention 33 547-561.

[11] C W J Granger (1969) Investigating causal relations by econometric models and cross spectral methods. Econometrica 37, 424-438

[12] J K Bae and J Kim (2010) Integration of heterogeneous models to predict consumer behaviour. Expert Systems with Applications 37(3), 1821-1826.

[13] F Akhter, D Hobbs and Z Maamar (2005) A fuzzy logic-based system for assessing the level of business-to-consumer (B2C) trust in electronic commerce. Expert Systems with Applications. 28, 623-628.

[14] K Miloslova & K Jiri (2008) Air Quality Modelling by Decision Trees in the Czech Republic Locality. 8th WSEAS International Conference on Applied Informatics and Communications (AIC'08) Rhodes, Greece, 196-201.

[15] C Huang, Y Lin and C Lin (2008) Implementation of Classifiers for Choosing Insurance Policy Using Decision Trees: A Case Study. Wseas Transactions on Computing, Vol. 7(10), 1679-1689.

[16] N Sarlija, M Bensic and M Zekic-Suzac (2006) Modelling customer revolving credit scoring using logistic regression, survival analysis and neural networks. Proceedings of the 7th WSEAS International Conference on Neural Networks, Cavtat, Croatia, 164-169.

[17] L Yang and C Chiu (2006) Knowledge discovery on consumer churn prediction. Proceedings of the 10th WSEAS International Conference on Applied Mathematics, Dallas, Texas, USA, November 1-3, 523-528.

[18] K Orozalieva (2010) Impact of Globalization on Socio-Economic and Political Development of the Central Asian Countries. A thesis submitted in partial fulfilment of the requirements for the degree of Master of Arts, Department of Government and International Affairs, College of Arts and Sciences, University of South Florida.

[19] Y Fujii, A Miura, J Tsukamoto et al. (1992) On-line expert system for power distribution system control. International Journal of Electrical Power & Energy Systems. Vol. 14, 1, 45-53.

[20] J M Hilb (2005) A review of SPSS, Part 3. The American Statistician. Vol. 59(2). 185-186.

[21] S C Lemon, J Roy, M A Clark, P D Friedmann, W Rakowski (2003) Source Classification and Regression Tree Analysis in Public Health: Methodological Review and Comparison with Logistic Regression Authors. Annals of Behavioural Medicine, vol. 26(3), 172-181.

[22] L Che-Chern, Y Hung-Jen; K Lung-Hsing (2009) Behaviour analysis of internet survey completion using decision trees: An exploratory study. Online Information Review, Vol. 33(1). 117-134.

[23] Y Xue-Cheng, J Wu, Z Xiao-Hang, L Ting-Jie (2008) Using decision tree and association rules to predict cross selling opportunities. Machine Learning and Cybernetics, 2008 International Conference on Volume: 3. 1807-1811.

[24] S Tsang, K Ben, K Yip, H Wai-Shing, L D Sau (2011) Decision Trees for Uncertain Data. Knowledge and Data Engineering, IEEE Transactions. Vol. 23(1). 64-78.

[25] X Mei-Ping, Z Wei-Ya (2010) The analysis of customers' satisfaction degree based on decision tree model. Fuzzy Systems and Knowledge Discovery (FSKD). Vol. 6. 2928 -2931.

[26] J Pereira-Moliner, Calver-Corttes E and J F Molina-Azorin (2011) Explaining the strategic group's firm performance relationship: A multilevel approach applied to small and medium sized hotel companies in Spain. Journal of Small Business Management. Vol. 49(3). 411-437.

R Basha received her PhD in April 1984 on Intelligent Information Systems from the University College of Cardiff, UK and worked at the Computer Science Department, University of Baghdad for several years where she was promoted to Associate Professor.
In Sept 1994, Dr. Basha joined Al alBayt University in Jordan where she served concurrently as Chairperson of Computer Science Department and Director of the Computer Centre. Later, she moved to the UAE as an Associate Professor at the University of Sharjah where she contributed to the Department of Computer Science for over 10 years. As part of her scholarly work, Dr. Basha has supervised 12 Masters' students and has over 20 publications in international and regional journals and acted as a referee on numerous international journals. She is currently with the College of Computer Information Technology, American University in the Emirates (AUE) Dubai International Academic City-Dubai UAE.

J Ameen received his PhD in May 1984 on Bayesian Dynamic Modelling and Forecasting from the University if Warwick, UK and worked for over seven years at Unilever Research, Port Sunlight, UK as a scientist with responsibilities for developing new models for market analysis and he is at his current position at the University of Glamorgan, South Wales UK since 1993 in which he has directed and supervised over 14 PhD and MSc students in various areas of time series, economics, banking, engineering and health care. Prof Jamal has widely published in international journals and acted as a referee on a number of international journals including The Journal of Forecasting, British Medical Journal, Journal of Advanced Nursing and The Journal of Construction and Economics. He is currently acting as Senior Advisor for the Ministry of Planning, Kurdistan Regional Government on the request of the former Prime Minister of Region.