AN ANALYSIS OF PREDICTING RETAIL CUSTOMER'S BEHAVIOUR USING ARTIFICIAL NEURAL NETWORKS

Dr.M.Kishore Babu, Professor & Dean KL University, Vijayawada, Andhra Pradesh, India.
B.Naveen Kumar, Research Scholar KL University, Vijayawada, Andhra Pradesh, India.
G.Anil Kumar, Research Scholar KL University, Vijayawada, Andhra Pradesh, India.
B.Radha Krishna, Research Scholar Osmania University, Hyderabad, Telangana, India.

Abstract
The focus of this research is to examine how different demographic features affect consumer behaviour in brick-and-mortar stores via the use of intelligent modelling methodologies. In order to foretell how retail customers would behave given their demographic information, this research use artificial neural networks (ANNs). This investigation was conducted in the UK. Individuals' demographic information such as age, income, family size, favourite retail business, frequency of monthly shopping trips, and average monthly spending are all taken into consideration. Artificial Neural Networks, or ANNs for short, are designed to mimic the communication patterns of human brain cells by simulating the interactions between computer systems. Quantifying the relationships between neurons in an artificial network requires a learning technique based on statistical concepts (ANN). While Artificial Neural Networks (ANN) have the potential to be used in the field of management, the field of engineering has seen the biggest use of ANN until recently. There has been a lot of research and development into applying ANNs in corporate settings, particularly in areas that were previously designated for multivariate statistical analysis. The reason for this transition is that artificial neural networks (ANNs) may outperform multivariate statistical analysis in certain scenarios. As a result of this transition, more attention is being paid to the areas that were formerly saved for multivariate statistical analysis. The accuracy of the model was tested by applying it to consumer survey data, and the findings confirmed the model's validity. The research for this study was place over the period of four months in 2019.

Keywords: Consumer Behaviour, Artificial Neural Networks, Retail

Introduction
Consumer behaviour is a study of individuals, both independently and in groups, with respect to their purchase and non-purchase behaviour. It covers what they buy and don't buy and what influences their purchase behaviour with respect when, how, why and where they prefer to buy. Essence of Economics, Sociology and Psychology can be found in Consumer behaviour studies. Consumer behaviour not only studies individual consumer's demographical characteristics like income, age, gender etc. but also behavioral characteristics like attitude, personality, motivation, learning etc. Further, consumer behaviour also studies the group behaviour of individuals i.e. how individuals behave when they are in groups like family, friends, and society.

In recent times, marketers have focused much on consumer behaviour studies as with passage of time, consumer behaviour has become a complex mix of several variables. This complexity of variable mix can be attributed largely to the wide choice of goods and services. Today, every organisation wants a share in the consumer's pocket and are offering a vast range of products/features/services using marketing mix.
It is becoming more popular for enterprises to utilise quantitative analysis for the aim of developing forecasts in fields like business and marketing. If a business takes the time to get to know its customers and make accurate predictions about the preferences they will have in the future, it can greatly improve its ability to market its wares to specific groups of consumers, reduce the amount of money it spends on marketing, increase the number of sales it makes, and significantly increase its overall profitability (Christopher et. al.). The ability of ANNs to more closely mimic human decision-making processes is what has sparked so much interest in them. This is because ANN can

more accurately and precisely represent the complexity of the consumer decision-making process. This is due to ANNs' superiority over competing methods in accurately representing nuanced customer behaviour. Models of artificial neural networks (also known as ANNs) are built with the human brain's structure in mind. They use a data-driven intelligence method to achieve their aim of conquering challenging problems in areas including pattern recognition, image processing, voice recognition, forecasting, and many more applications of data mining. In this research project, we aimed to apply the MLP model to foretell shoppers' behaviours at a store. The model consists of a single hidden layer that collects data from the other five variables all at once.

Review of Literature

Several fields, including medical, mathematics, psychology, statistics, and computer science, have discovered uses for ANN. Manufacturing are only few of the recent commercial domains that benefit from ANN applications (Vellido et al., 1999a). To determine whether method was superior for classifying consumers into risk-taking and safe categories, Dasgupta et al. (1994) tested ANN models against discriminant analysis and logistic regression.

The computational operations of the brain constitute a highly advanced, nonlinear, parallel information processing system. The brain is able to conduct some computations far more quickly and accurately than the finest computers do because of the way it arranges its components. An artificial neural network is a simplified computer simulation of the brain's neural circuits.

The number of neurons in the human brain is believed to be 1011. There are probably about 1015 connections between these neurons. An ANN, like a brain, is made up of artificial neurons that employ a processing element set that is similar to that of real neurons and their interconnections. The network containing these nodes is capable of recognising patterns in data.

Pattern recognition, robotics, finance, aerospace, medicine, telecommunications, manufacturing, speech, control systems, and security identification using classification, function approximation, prediction, and clustering are just some of the areas where neural networks can be trained to solve problems that are challenging for traditional computers or humans. ANN's superior ability in mining non-linear data gives it an edge over conventional models.

Bruce Grey first proposed the concept of neural marketing in 1992. Using a Neural Network, he developed a marketing model that takes into account a wide range of variables, including sales, GDP growth, survey responses, firm size, and industry.

To study how shoppers act, Luiz A. Moutinho (and coworkers) built an ANN model. Customers with disposable income to take advantage of the variety and quality on offer tend to be the most happy and spend the most money at the supermarket, according to an analysis of many key characteristics. Dr. Naveen Prasadula demonstrated that Neural Networks outperformed in analysing competitive market structure and anticipating asymmetric cross-elasticities. In 2005, Christopher Gan et al. employed a Multilayer Feedforward Network and a probabilistic Neural Network to examine customers' preferences between online and traditional banking.

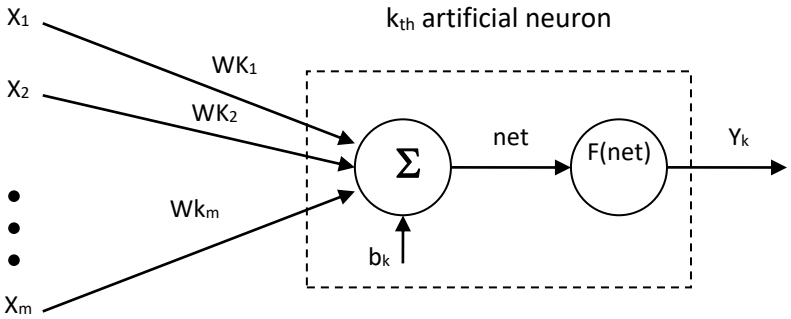
To improve the Customer service, a survey was conducted in a shopping mall in China, about their beliefs and evaluation of goods purchased. Objective and subjective attributes of consumer products were analyzed to know about the consumer’s attitude and perceptions. The Neural Network outputs were compared with Correlation and Regression Trees (CART) model to assess the significance of ANN. The ANN model surpassed the CART model with 89.5% accuracy, where as 82.6% accuracy was achieved through CART model.

The Composition of Synthetic Neural Networks

The processing power of neural networks is very sensitive to the organisation of neural processing units and the quality of the connections between them. “Number of layers and number

ISSN: 2278-4632
Vol-11 Issue-12 No.01 December 2021

of nodes in each layer are often used to describe the structure (Minsky –



1954).

Artificial neuron model; Figure 1.

The model of the neuron given above also includes an externally applied bias, denoted by b_k. The bias has the effect of increasing or lowering the net input of the activation function, depending on whether it is positive or negative". The net summation will be

$$\mathbf{net_k = x_1w_{k1} + x_1w_{k1} + x_2w_{k2} + \ldots + x_mw_{km} = \Sigma x_iw_{ki}}$$

Finally, an artificial neuron computes the output y_k as a certain function of net_k value:

$$\mathbf{y_k = f(net_k)}$$

Multilayer Perceptron Network Architecture

Half of all neural network business application studies, according to the most recent study, use multilayered neural networks using the back propagation learning method (Wong et al., 2000). This is one of the situations in which you could require their help. As they need a dataset to get familiar with the data and identify patterns within it, they fall under the category of supervised learning. In order for them to learn, they need to have access to the dataset. This makes it crucial for them to learn from a human teacher.

A complex hierarchy of processors ultimately determines the network's output. Each of these processors does some calculation on the data it receives and passes the results on to the layer above it. Each individual input signal is then subjected to the threshold value before being combined with the other input signals by the processing unit. The whole input signal is sent via an activation function to determine the actual processing unit output, which is then used as the input to a subsequent layer of processing units in a multilayer network. This procedure will be repeated until the processor's true output is determined. This process will continue until the precise output of the CPU is determined.

As can be seen in Figure 2, the "MLP design includes not only an input layer and an output layer of neurons, but also a hidden layer or layers of neurons in between. The inputs X are transferred through the network through the connecting weights to the hidden layer of neurons, and then to the output layer, in a left-to-right fashion". Wji represents the weights that send information from node I to hidden node j, and Vkj represents the weights that send information from node j to node k as output.

ISSN: 2278-4632
Vol-11 Issue-12 No.01 December 2021

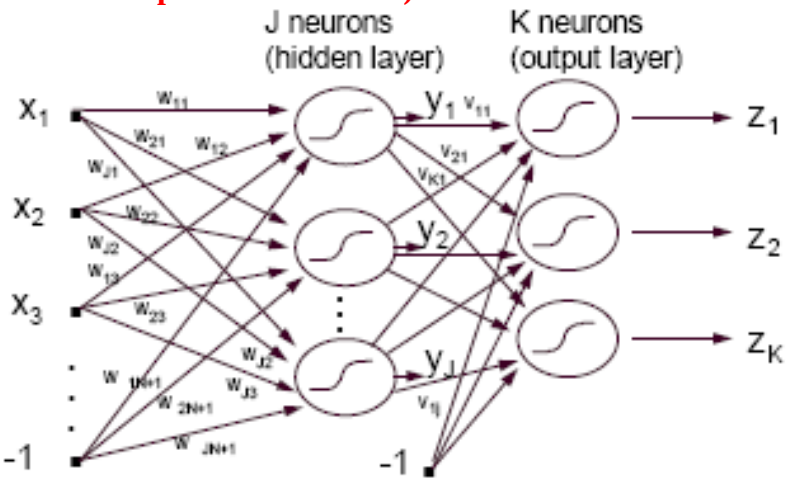


Fig 2: Architecture of Multilayer Perceptron Network

Neurons determine their output in relation to the amount of stimulus they get. The net input, which is calculated as the weighted total of the neuron's inputs, determines the size of the output, which is based on a sigmoidal function. To be more precise, this pertains to the jth undisclosed neuron.

$$net_j^h = \sum_{i=1}^{N+1} W_{ji} x_i \; and \; y_j = f(net_j^h)$$

while k-th neuron output $net_k^o = \sum_{j=1}^{J+1} V_{kj} y_j \; and \; z_k = f(net_k^o)$

The well-known continuously differentiable logistic function f(net) is the typical example of a sigmoidal function. If the total input is less than a threshold value (v), the Threshold Function returns 0, and if it's larger than or equal to v, it returns 1.

$$f(net) = \frac{1}{1 + e^{-\lambda net}}$$

where is a gradient-controlling parameter; the function's gradient need only be growing monotonically and differentiable within the range [0,1]. The expected response from each neuron is often a continuous value when dealing with prediction difficulties. The network's weights are adjusted to compensate for the inaccuracy, and the next pattern is shown. The network then computes its output and checks it against the measured result and the estimated margin of error.

Back-propagation is the name given to the corresponding learning algorithm. The final mapping of activations of output neurons into network outputs is performed by the threshold function momentum, which makes use of the previous increase to the weight to speed up and stabilise the convergence.

Data Collection and Research Methodology

A questionnaire was used to compile the data. Questions in the survey focused on respondents' ages, incomes, family sizes, shopping habits, average purchases, and preferred stores.The sample respondents consisted of customers at three hypermarkets; BigBaazar (Abids), Reliance Mart (Toli Chowki) and Spencer's (Musheerabad). The sample consisted of total 487 respondents. The poll was administered throughout the store's outside areas, including the food courts and sit-outs.

The Neural Networks method began with the creation of a model consisting of five input nodes and a single output node.

The input nodes are:

- Age

ISSN: 2278-4632
Vol-11 Issue-12 No.01 December 2021

- Income
- Family size
- Frequency of visits
- Average amount spent on each shopping visit

The output node is:

- Store choice

a Multilayer Perceptron Network model was built with the following architecture.

Table 1: Model Architecture `

"Input Layer	Covariates	1	No of People
		2	Age
		3	Income
		4	How Often you Shop
		5	Avg Amount Spent
	Number of Units[a]		5
	Rescaling Method for Covariates		Normalized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		4
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Name of the Store
	Number of Units		3
	Activation Function		Softmax
	Error Function		Cross-entropy"

a. Excluding the bias unit

Results

The purpose of this research was to see whether, by evaluating demographic information, an MLP neural network might aid in predicting a customer's preferred retail location. The datasets used to train the ANN model are listed in Table 2.

Overview of Case Management (Table 2)

"		N	Percent
Sample	Training	296	60.8%
	Testing	88	18.1%
	Holdout	103	21.1%
Valid		487	100.0%
Excluded		0	
Total		487	

The network diagram that SPSS used to predict course outcome (Var01=Big Bazaar, Var02=Spencer's and Var03=Reliance Mart) from 5 demographic characteristics of customers as shown in Figure 3. The diagram shows the 5 input nodes, the 4 hidden nodes and the one output node representing the retail outlets.

ISSN: 2278-4632
Vol-11 Issue-12 No.01 December 2021

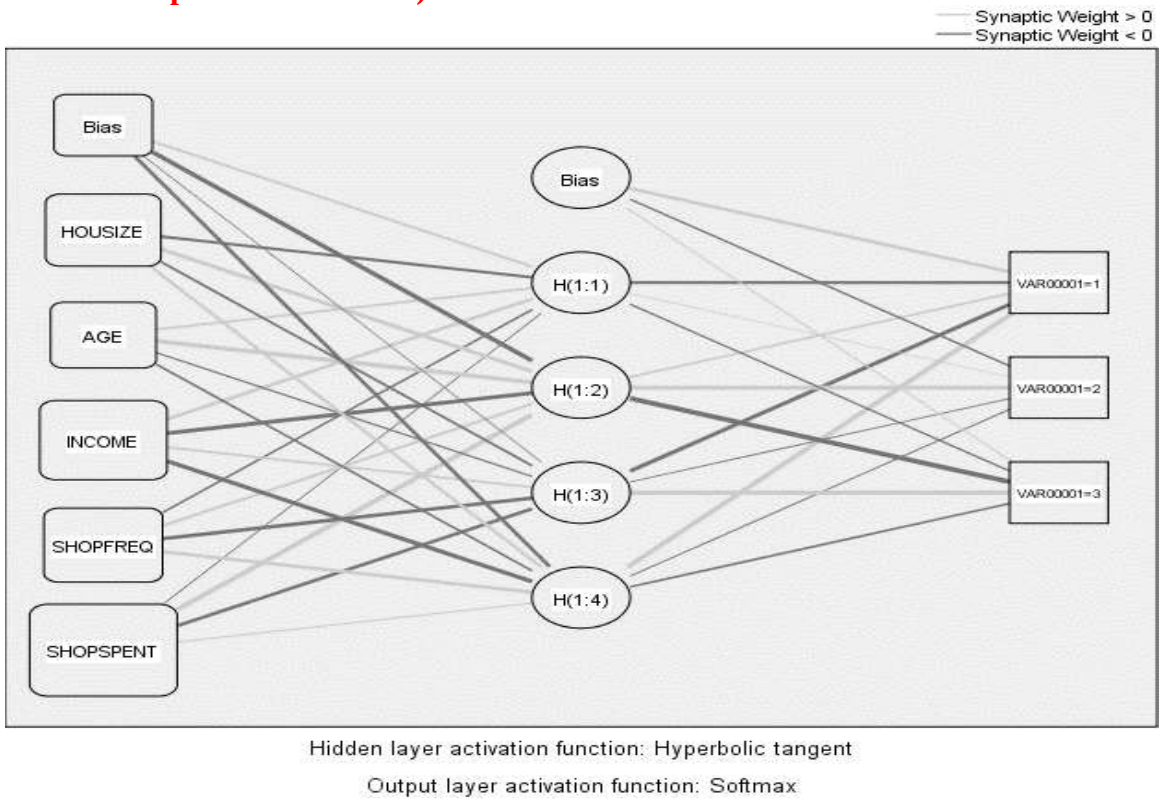


Figure 3: Network Diagram

The model summary, shown in Table 3, provides information related to the results of training, testing and holdout sample. The incorrect predictions based on training, testing and holdout samples are 47.6%, 39.8% and 52.4% respectively. The learning procedure was performed for 1 consecutive steps with no decrease in error function was attained from the testing sample.

Table 3: Model Summary

Training	Cross Entropy Error	283.578
	Percent Incorrect Predictions	47.6%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.08
Testing	Cross Entropy Error	80.790
	Percent Incorrect Predictions	39.8%
Holdout	Percent Incorrect Predictions	52.4%

Dependent Variable: Name of the Store
a. Error computations are based on the testing sample.

Table 4 displays the synaptic weights between which have been calculated using only the data of the training dataset

Table 4: Parameter Estimates

		Predicted						
		Hidden Layer 1				Output Layer		
Predictor		H(1:1)	H(1:2)	H(1:3)	H(1:4)	[VAR00001=1]	[VAR00001=2]	[VAR00001=3]
Input Layer	(Bias)	.299	-.653	-.029	-.376			
	HOUSIZE	-.354	.553	-.221	.359			

ISSN: 2278-4632
Vol-11 Issue-12 No.01 December 2021

	AGE	.248	.701	-.120	-.206			
	INCOME	.465	-.563	.246	-.638			
	SHOPFREQ	-.197	.317	-.525	.518			
	SHOPSPENT	-.071	.984	-.463	.106			
Hidden Layer 1	(Bias)					.485	-.190	.082
	H(1:1)					-.368	.063	-.144
	H(1:2)					.221	.694	-1.435
	H(1:3)					-.537	-.078	.929
	H(1:4)					.933	-.079	-.209”

Table 5 is a categorization of the dependent categorical variable Store selection result by group and as a whole. If the anticipated probability is larger than 0.5 for each observation, then the expected output is considered successful. The MLP network accurately classified 155/296 customers in the training sample, and 53/88 customers in the testing sample.Overall, the training examples were accurately categorised in 52.4% of all instances. Big Bazaar has a significantly more optimistic forecast than either Spencer's or Reliance Mart, as seen in the table. For Big Bazaar, the MLP network had a prediction accuracy of about 75%, whereas for Reliance Mart it was around 50%.MLP network failed completely in case of Spencer’s. This shows high consistency in the patterns for the customers for Big Bazaar where as there is high inconsistency in the patterns for the customers of Spencer’s and Reliance Mart.

Table 5: Classification

		Predicted			
“Sample	Observed	Big Bizaar	Spencers	Reliance Mart	Percent Correct
Training	Big Bazaar	96	14	21	73.3%
	Spencer’s	57	19	15	20.9%
	Reliance Mart	26	8	40	54.1%
	Overall Percent	60.5%	13.9%	25.7%	52.4%
Testing	Big Bazaar	32	5	2	82.1%
	Spencer’s	17	8	5	26.7%
	Reliance Mart	6	0	13	68.4%
	Overall Percent	62.5%	14.8%	22.7%	60.2%
Holdout	Big Bazaar	38	3	10	74.5%
	Spencer’s	23	5	6	14.7%
	Reliance Mart	10	2	6	33.3%
	Overall Percent	68.9%	9.7%	21.4%	47.6%

Dependent Variable: Name of the Store

The proportional and normalised significance of each independent variable in the ANN model is shown in Table 6.

Table 6: Independent Variable Importance

	Importance	Normalized Importance
No of People	.180	67.5%
Age	.153	57.3%
Income	.212	79.4%
How Often you Shop	.189	70.8%
Avg Amount Spent	.267	100.0%”

Conclusion

This study set out to see whether artificial neural networks might be used to reliably predict customer purchasing behaviour given knowledge of their demographic characteristics.The classifier has a very high success rate of 75% when asked to categorise Big Bazaar clients into groups with high and low probabilities of success and failure. The investigations also uncovered the most influential factors in customers' choice of brands to support. This makes it harder to find a solution to the problem. This is due to the fact that nonlinear functions are only one kind of neuron in a neural network. This is due to the fact that analysing dependent variables is often more challenging than analysing independent ones. The reason for this is because no one approach can simultaneously achieve both objectives. This is due to the fact that there are currently no standard procedures for handling non-linear processes. This is a key drawback of current methods that must be addressed. However, it's vital to note that training neural networks to acquire new knowledge might be a lengthy process.While the literature suggests that ANN may be utilised effectively to foresee the actions of customers, employees, and even students, more research is required to corroborate these findings using larger and more diverse samples. This much is accurate, however it will need additional study to corroborate the findings with larger and more representative samples. This is where we are currently; however, further study is required with bigger and more representative samples to confirm these results. This will become very important at some point in the near future.

References

1. Blackwell, Miniard and Engel (2006). Consumer Behaviour (10th Ed.). Thomson Learning.
2. Bruce Grey Tedesco, “Neural Marketing: Artificial Intelligence Neural Networks in Measuring Consumer Expectations”, Grey Associates, Inc. Chicago IL, 1992
3. Christopher Gan, Visit Limsombunchai, Mike Clemes and Amy Weng , “Consumer Choice prediction: Artificial Neural Networks Versus logistics Models”, Commerce Division, Discussion Paper No. 104, ISSN 1174-5045 , ISBN 1-877176-81-8, 2005
4. Dasgupta, Bhaskar & Siegelmann, Hava & Sontag, Eduardo. (1994). On a Learnability Question Associated to Neural Networks with Continuous Activations.
5. Fish, K.E., Barnes, J.H., & Aiken, M.W.(1995). Artificial neural networks — A new methodology for industrial market segmentation. Industrial Marketing Management,24, 431–438. Harry B. Burke, MD., Ph.D., David B. Rosen, Ph.D., Philip H. Goodman, “Comparing Artificial Neural Networks to Other Statistical Methods for Medical Outcome Prediction” IEEE/1994/(Page 2213 – 2216).
6. Hu, M.Y., M. Shanker, and M.S. Hung (1999), “Estimation of posterior probabilities of consumer situational choices with neural network classifiers”, International Journal of Research in Marketing, 16(4), 307-317.
7. Jaesoo Kim, Heejune Ahn , “A New Perspective for Neural Networks: Application to a Marketing Management Problem”, Journal of Information Science and Engineering 25, 1605-1616(2009)
8. J. Scott Armstrong (1991). "Prediction of Consumer Behavior by Experts and Novices". Journal of Consumer Research Inc.. 251–256.
9. Dr Naveen Prasadula(2021) : A Review of Literature on An Analysis of Predicting Retail Customer’s Behaviour using Artificial Neural Networks
10. Kate smith, Jatinder Gupta , “Neural Networks in Business: Techniques and Applications”, Idea group and Information Science Publishing.
11. Luiz A, Mountinho, Fiona M. Davies, Mark M. H. Goode, Emmanuel Ogbonna, “ Critical facors in consumer supermarket shopping behavior: A Neural Network Approach”, Journal of consumer Behaviour, Vol 1, Issue1,Page 35-49,2001.
12. T.S. Gruca, B.R. Klemz , “ Using neural networks to identify competitive market structures from aggregate market response data”, Science Direct, Omega, Vol. 26, Issue 4, Page 549.
13. Vellido, Alfredo, Lisboa, P.j.g, Vaughan.J, "Neural networks in business: A survey of applications (1992-1998)", Expert Systems With Applications - ESWA Vol 17, pg. 51-70, 1999

