

Customer Retention in Banking Sector using Decision Tree-Neuro Based System

Tamilmani.G, Rajathi.K



Abstract: All the bank marketing campaigns mostly deals with large amount of data. when they need to deal with huge electronic data of customers, then it is very difficult to analyze the data manually or by human analyst. Here comes the picture of data mining techniques to deal with the large amount of data and to come up with useful data which helps in decision making process. All the data mining techniques helps in analyzing the data. some of the techniques that can be used for this bank marketing campaigns are naive bayes, logistics regression technique and Decision tree model technique etc. among all these techniques decision Tree model gives the best solution in analyzing the human decisions. Artificial neural networks is a learning algorithm which learns from multiple individual decisions and their judgements, thus aggregates and generalizes the customers decision making knowledge.

Keyword: Decision Tree Model Technique, Artificial Neural Network, logistics regression.

I. INTRODUCTION

All the banks contain large amount of data regarding their customer details. Which is used to directly attack the customers with this information by phone calls or messages or emails when they need to promote a new product or new service is called **Direct marketing**. In fact, direct marketing is the best strategy for the bank people or insurance companies when they need to advertise the new product.

II. DESIGN TECHNIQUES

A. Methodology

Here, the decision trees and neural networks are combined to form a hybrid methodology called as **Decision Tree-Neuro based system**.

B. Analysis of the method

Every methodology has their own advantages and disadvantages. Where decision tree can do few things, which can't by neural networks, as well neural networks can do few which can't by decision tree technique. For example, artificial neural networks can efficiently deal with training

data where decision trees can't, decision trees are fast in learning where the neural networks are little bit slow process compare to decision trees. So, the combination of both leads to the best outcome.

III. DECISION TREE NEURO BASED SYSTEM DESIGN

In this decision tree neuro-based system, this consist of two parts as shown in the fig (3). The first part consists of the decision tree, how the root node is classified into internal nodes and Internal nodes splits to leaf nodes or terminal nodes. The leaf nodes are the final outcome of the decision trees, where the output of the decision tree becomes an input to the neural networks. And starts the further processing which results in the high level of accuracy in decision making. Basically, human beings are the one who takes the decision based on some suggestions from the people or some by some actions. So, this condition is deployed in developing the algorithm. Here the decisions of the decision tree are not the final outcome, these are just only input to the neural networks. Where the neural networks process the data based on how the decision tree is performed, and the final decision of the problem is done by the neural networks. The process of designing a neural network is a logical process. Here the procedure is not done by single time, this has to be done multiple times like going back to the first step again and again so the hidden layer processes the data several times based on the threshold value. the process is done as many times as possible until the error rate is minimized. Here the bank rules are used as their parameters for granting their services. The nodes will be divided based on the bank conditions for each and every service. Once all the test conditions are done the classification of the nodes will be stopped and the outcome is given as input to the neural networks to improve the efficiency of the decision making. The number of attributes that are given as input to the neural network is based on the outcomes found by decision tree.

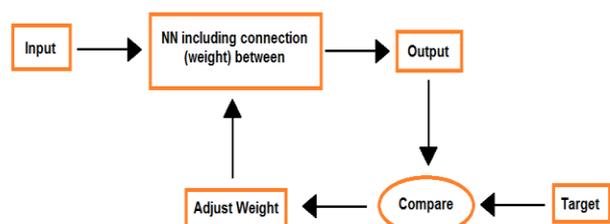


Fig 1: Adjusting the neural network to minimize the error rate by comparing with target.

Revised Manuscript Received on February 05, 2020.

* Correspondence Author

Tamilmani.G*, Assistant Professor, Computer Science and Engineering, Vel tech Rangarajan Dr.Sagunthala R & D Institute of Technology, Avadi, Chennai. Email: tamilmanig@veltech.edu.in

Rajathi.K, Associate Professor, Computer Science and Engineering, Vel tech Rangarajan Dr.Sagunthala R & D Institute of Technology, Avadi, Chennai. Email: rajathi@veltech.edu.in

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)



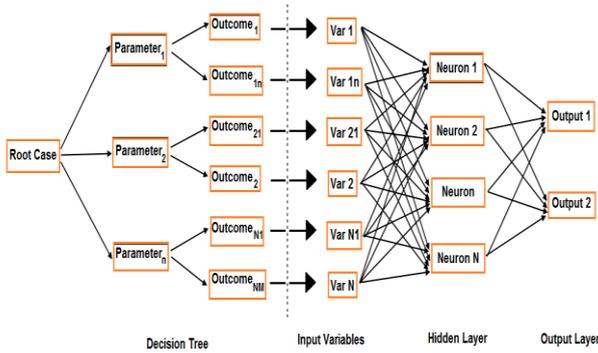


Fig 2: Decision tree neuro-based system

IV. DECISION TREE ALGORITHM

Decision tree algorithm has more influence in machine learning, where decision tree comes under supervised machine learning technique, which will be like a tree structure with a definite hierarchy. Decision tree consist of links and nodes. Each link carries the decision of a particular node and leaf node represents the outcome of the parent node decision. **It is the best technique to analyze the people decisions.** This algorithm is commonly used for classification and regression problems. All the machine learning concepts will deal with the linear data, where the decision tree algorithm has the capability to deal with the non-linear data.

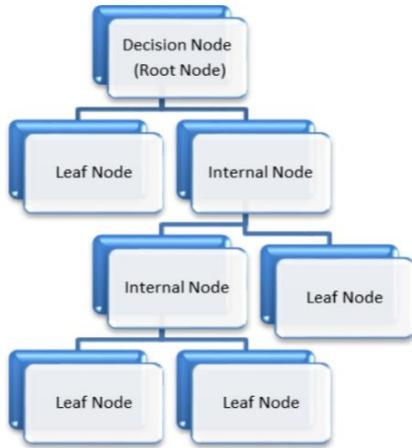


Fig 3: Decision Tree

As shown in the above figure, Decision tree contains **Root Node, Internal Nodes, Leaf/Terminal Nodes and Branches/Links.**

Root Node is the initial point of the Decision Tree where actual splitting will start from the Root node. **Internal Node** each internal node has a (predictor variable) or a decision value, which leads to get the final outcome of the give data set or problem. **Leaf Nodes/Terminal Nodes** is also called as terminating point, because this is the end point of the decision tree where the splitting will stop's and the output of the leaf node is the final outcome/output. **Branches** are used to connect the nodes; this is represented as Arrows/links. Each link says weather the parent node decision is yes or no.

V. HOW THE DECISION TREE ALGORITHM WORKS

Step 1: At every level select the best feature as a Node. i.e, Root node.

Step 2: Now classify the node into internal nodes which helps in getting the final outcome.

Step 3: Repeat the steps 1 and 2 until the data is properly classified that is all the test conditions are completed.

Here the first step is to select the best feature as a test condition, so what does the best feature mean, the best feature or (predictor variable) is the one that splits the data set into different classes more efficiently, which leads to the best outcome.

How to decide whether it is a best feature or not, the best feature can be measured in terms of

1. Entropy
2. Information Gain (IG).

Entropy is used to measure the ambiguity present in the data, (i.e.) it will decide how efficiently the decision tree can classify the data or split the data.

Equation: $Entropy = -\sum p(x) \log(x)$

Information Gain (IG) plays major role in building decision tree, where the information Gain indicates, how much each node or feature can give the information about final output. The variable, which has the highest IG will be selected as a root node, and the actual splitting will start from root node.

Equation: $Entropy(parent) - [weighted\ average] * entropy(children).$

VI. NEURAL NETWORKS MODELLING

Artificial neural networks are a method where the information is processed like how the human brain process the data. neural networks learn by training data with an algorithm. Once the training of data is done the neural networks process the data by comparing the data with the past experience. Here every neuron will get the sum of inputs, so each neuron processes the given inputs. The output of neural networks depends on the combination of all the neurons. So, all the neurons need to process the data in an efficient manner. This neural network is used in many applications like predictions, clustering, classification, categorization etc., the outcome of the processed neurons is calculated with the following equation:

$S_j^l = g(\sum_{j=1}^m s_j^{l-1} w_j^l - \theta_j^l)$

Where S_j^l is the output of the neuron in the layer l.

s_j^{l-1} is the output of the neuron j in the layer l-1.

w_j^l is the weight of each neuron in layer l

θ_j^l is the threshold value of the neuro in layer l

Here every network is assigned with a weight, and all the weights are inter connected. The algorithm contains the set of inputs and the expected output. so the neural network undergoes supervised learning mechanism,



so that the algorithm computes the data with the given inputs and the actual output is compared with the expected output. So the difference or error between the

actual output and the expected output is calculate. And the data is processed several times in hidden layer so that the difference between the actual output and expected output is minimized. The main goal is to reduce the error rate till the neural networks learns the training data. Initially the training starts with the random weights, and finally the weights are modified during the process to reduce the error rate.

A. Data for the system

Bank marketing data set is used from the university of California to test the efficiency of the decision tree algorithm and neural networks that is decision tree neuro-based system. The data set is completely related to bank marketing campaigns of a Portuguese banking institution. Here the marketing is done through phone calls. Frequently the phone calls are done to know whether the service or product is subscribed or not. The data set consist of (44122) no of samples with 15 attributes without missing any values. The data set is classified into two categories 1. Nominal attributes 2. Numerical attributes. Numerical attributes represents the data set like age, pday, previous, campaign, balance, day,

duration etc., binary data set represents the values of yes or no. Each attribute consists of number of classes that is represented in column attributes. For example a job attribute has number of classes like (unemployed, student, teacher, software engineer, self-employee, business, etc,.). A girl or women is divorced means there exists again two classes like whether she is divorced or widowed. In the bank direct marketing telephonic calls are done to the clients and all the details are gathered, tabular column is maintained for the client data, like how many days before the person is contacted. What is the status of client, weather in previous campaign, we contacted the client or not etc,. There are two machine learning techniques to analyze and predict the data 1. Supervise learning technique and 2. Unsupervised learning technique. Supervised learning technique requires two input data first one is predictor variable that is (Independent data) and the second one is target data that is (dependent data). So, the algorithm compared the predictor variable with the target data and process the data. Best technique in supervised learning is decision tree technique, actually supervised learning is the best one to predict and analyze the nonlinear data. Unsupervised learning is used to analyze the clustered data, statistical measures etc,

Table - I: Attribute description Experimental Results

#	Attributes	Kind	Type	Attributes Illustration	Domain
1	Age	Numeric	Range		18..95
2	Education	Categorical	Set	('unknown','secondary','primary','tertiary')	
3	Job	Categorical	Set	('admin.','unknown','unemployed','management','housemaid','entrepreneur','student','blue-collar','self-employed','retired','technician','services')	
4	Marital	Categorical	Set	marital status ('married','divorced','single'; note: 'divorced' means divorced or widowed)	
5	Balance	Numeric	Range	average yearly balance, in euros	8019:102127
6	Default	Binary (Categorical)	Flag	has credit in default? (binary: 'yes','no')	
7	Contact	Categorical	Set	contact communication type (categorical: 'unknown','telephone','cellular')	
8	Housing	Binary (Categorical)	Flag	has housing loan? (binary: 'yes','no')	
9	Loan	Binary (Categorical)	Flag	has personal loan? (binary: 'yes','no') # related with the last contact of the current campaign	
10	Duration	Numeric	Range	last contact duration, in seconds	0..49
11	Month	Categorical	Set	last contact month of year (categorical: 'jan', 'feb', 'mar', ..., 'nov', 'dec')	
12	Day	Numeric	Range	last contact day of the month	1:31
13	Campaign	Numeric	Range	number of contacts performed during this campaign and for this client (includes last contact)	
14	Pdays	Numeric	Range	number of days that passed by after the client was last contacted from a previous campaign (-1 means client was not previously contacted)	1:87
15	Poutcome	Categorical	Set	outcome of the previous marketing campaign (categorical: 'unknown','other','failure','success')	
16	Previous	Numeric	Range	number of contacts performed before this campaign and for this client	0:275
17	Output	Binary (Categorical)	Flag	output variable (desired output): y - has the client subscribed a term deposit? (binary: 'yes','no')	

The efficiency of each prediction model is calculated in terms of Accuracy, sensitivity and specificity. These three statistics are based on **True Positives (TP)** this indicates the number of correct predictions that is true. **True Negatives (TN)** this indicates the number of true predictions that are false. **False Positives (FP)** this indicates the number of false predictions that are true. **False Negatives (FN)** this indicates the number of false predictions that are false. Accuracy, sensitivity and specificity can be defined with the following equations:

VII. RESULTS

Accuracy : (TP + TN) /N

Accuracy can be calculated by adding true positives and true negatives and divide the resulted value with total number of attributes.

Sensitivity: TP / (TP + FN)

Sensitivity can be calculated by dividing the true positive value with the added result of true positives and false negatives.

Specificity: TN / (TN+FP)

Specificity can be calculated by dividing the true negative value with the added result of True negatives and false positives.



Table -II: Comparison of DT vs NN vs DTNN

Model	Training Data			Testing Data		
	Desired output	Yes	No	Desired Output	Yes	No
DT	Yes	TP = 2012	FP = 2013	Yes	TP = 869	FP = 1000
	No	FN = 1083	TN = 31,123	No	FN = 630	TN = 22,635
NN	Yes	TP = 1193	FP = 1018	Yes	TP = 789	FP = 968
	No	FN = 983	TN = 23,543	No	FN = 589	TN = 11,983
DTNN	Yes	TP = 2258	FP = 1354	Yes	TP = 740	FP = 837
	No	FN = 684	TN = 28,193	No	FN = 512	TN = 11,321

Table-III: Accuracy, Sensitivity and Specificity comparison of models

Model	Partition	Accuracy	Sensitivity	Specificity
DT	Training	89.30%	58.90%	87.60%
	Testing	88.60%	61.40%	89.30%
NN	Training	90.54%	71.87%	90.87%
	Testing	91.52%	75.34%	91.13%
DTNN	Training	92.76%	81.64%	92.63%
	Testing	93.78%	85.67%	95.67%

The system is designed in a way that there exist three different cases first the performance of DT (Decision Tree) alone calculated and next the performance of neural networks (NN) alone calculated and finally the combination or hybrid of decision tree and neural networks that is DTNN (decision tree neuro-based system) alone calculated. So, from the graph it shows DTNN has high performance compare to the

DT and NN.

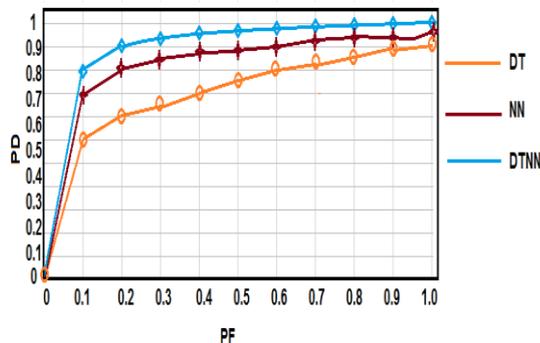


Fig 4: Graph Representation of the comparison of models

A. Percentile Training Data and Percentile Testing Data

Fig 4., shows the higher the line indicates the better performance. This above figure shows the DTNN (Decision tree based neural networks system) is the best in predicting the decisions of nonlinear data compare to the other techniques.

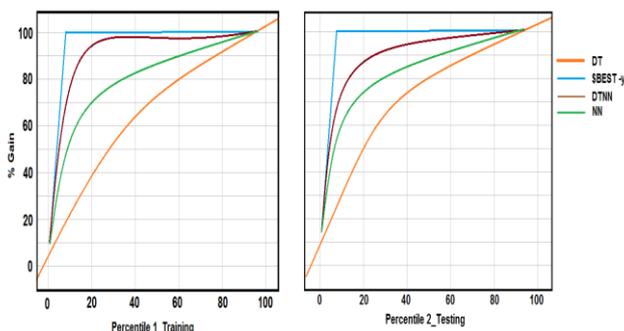


Fig 5: High Performance DTNN

VIII. CONCLUSION

Analyzing the human decisions is very complex. Now a days the decisions of human beings are changing dynamically (day to day life). So, to predict the human decisions, decision tree is the best algorithm compared to all other supervised and unsupervised algorithms. Here the combination of decision tree and neural networks that is decision tree neuro-based system provides 92% of accuracy in predicting the nonlinear data.

REFERENCES

1. Ihsan A. Kareem 1, Mehdi G. Duaimi 2, Improved Accuracy for Decision Tree Algorithm Based on Unsupervised Discretization, International Journal of Computer Science and Mobile Computing, Vol.3 Issue.6, June- 2014.
2. Linyuan Xia, Qiumei Huang , and Dongjin Wu, Decision Tree-Based Contextual Location Prediction from Mobile Device Logs,Hindawi Mobile Information Systems Volume 2018, Article ID 1852861.
3. Francisca Nonyelum Ogwueleka1, Sanjay Misra2, Ricardo Colomo-Palacios3, Luis Fernandez4 Neural Network and Classification Approach in Identifying Customer Behaviour in the Banking Sector:A Case Study of an International Bank, DOI: 10.1002/hfm.20398 .
4. Thitiman Damrongsakmethee 1*, Victor-Emil Neagoe 2, Data Mining and Machine Learning for Financial Analysis, DOI: 10.17485/ijst/2017/v10i39/119861.
5. Md. Rafiqul Islam1 and Md. Ahsan Habib2, A Data Mining Approach To Predict Prospective Business Sectors For Lending In Retail Banking Using Decision Tree, International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.5, No.2, March 2015.
6. Himani Sharma1, Sunil Kumar2,A Survey on Decision Tree Algorithms of Classification in DaInternational Journal of Science and Research (IJSR) ISSN (Online): 2319-7064Index Copernicus Value (2013)
7. Josip Mesarić1, and Dario Šebalj,Decision trees for predicting the academic success of students,Croatian Operational Research Review CRORR 7(2016).
8. Sujata S.Kharat, Vamshi Krishna, To Study Artificial Neural Networks in Data Mining and Its Method, Volume 3, Issue 7, July 2015 International Journal of Advance Research in Computer Science and Management Studies, ISSN: 2321-7782.

9. Tharaha, Rasika, Hybrid artificial neural network and decision tree algorithm for disease recognition and prediction in human blood cells, INSPEC Accession Number: 17558792 DOI: 10.1109/ICIECS.2017.8276107.
10. Brijesh Kumar Baradwaj, Saurabh Pal, Mining Educational Data to Analyze Students Performance, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.
11. Mr. Brijain R Patel, 2Mr. Kushik K Rana, A Survey on Decision Tree Algorithm For Classification, 2014 IJEDR | Volume 2, Issue 1 | ISSN: 2321-9939.
12. Bhaskar N. Patel, Satish G. Prajapati and Dr. Kamaljit I. Lakhtaria, Efficient Classification of Data Using Decision Tree, Bonfring International Journal of Data Mining, Vol. 2, No. 1, March 2012.

AUTHORS PROFILE



Tamilmani. G., received B.Tech degree in CSE from Avinashilingam Institute of home science and higher education for women, Coimbatore in 2011. M.Tech Degree in CSE from Avinashilingam Institute of home science and higher education for women, Coimbatore in 2013. She is currently working toward the Ph.D. degree at the Department of Computer Science and Engineering, Avinashilingam Institute of home science and higher education for women, Coimbatore. Her research interests include Data mining and Image Processing. Now currently working as Assistant Professor in Vel Tech Rangarajan Dr. Sagunthala R& D Institute of science and Technology in CSE, Avadi, Chennai



Rajathi. K., having 8 years of experience and currently working as Associate Professor in Vel Tech Rangarajan Dr. Sagunthala R& D Institute of science and Technology in CSE, Avadi, Chennai.