BANK CUSTOMER RETENTION PREDICTION AND CUSTOMER RANKING BASED ON DEEP NEURAL NETWORKS

Dr A.P.JAGADEESAN Ph.D
Head of Department Associate Professor/CSE
RVS College of Engineering Dindigul-624005

Indhuja
PG Scholar
RVS College of Engineering Dindigul-624005

***Abstract*: Retention of customers is a major concern in any industry. Customer churn is an important metrics that gives hard truth about the retention percentage of customers. A detailed study about the existing models for predicting the customer churn is made and a new model based on Artificial Neural Network is proposed to find the customer churn in banking domain. The proposed model is compared with the existing machine learning models. Logistic regression, Decision Tree and random forest mechanisms are the baseline models that are used for comparison, the performance metrics that were compared are accuracy, precision, recall and F1 score. It has been observed that the artificial neural network model performs well than the logistic regression model and decision tree model. But when the results are compared with the random forest model considerable difference is not noted. The proposed model differs from the existing models in a way that it can rank the customers in the order in which they would leave the organization.**

***Keywords*: Churn Prediction, Logistic regression, Decision Tree, random forest, Artificial Neural Network.**

I Introduction

Customers are the soul of any industry. In the competitive world, retaining the existing customers is a major concern. The customer retention process lies on the customer relationship management department of the organization [1]. It has also been observed that the retention of existing customers is less expensive than making new ones [2], the cost of which is several times higher [3]. The objective of this work is to analyse the various state of art models that are in use for churn prediction. The paper is organized as follows, the next section explains the various state of art models that are in practice, the third section explains the proposed Artificial neural network model. A detailed explanation of how the model is built and how the ranking of customers is made is given. Fourth section gives the results and comparison with the base line models. The last and the fifth section give the conclusion.

II Study of state of art models

This section explains the various works that have been done in order to predict the customer churn. It includes both the machine learning models and other models.

In addition to the conventional data used for predicting the customer churn, the authors of [4] have added data from the various sources. It includes the conversation of the customers through phone, the websites and products the customer has viewed, interactive voice data and other financial data. Linear regression model is used for predicting the customer churn. Though a good improvement is noticed with this model, the data that has been used in this is not commonly available at all the times.

In [5] the authors have also considered the textual data for predicting the customer churn. Around 23 fields were taken into account and applied in the convolutional neural network. When some other state of art works have used CNN for analysing time series data, the authors have claimed that textual data can also be efficiently analysed with CNN that could produce fruitful results. The authors have also specified that the designed CNN model extracts features in a better way and the relation between the churn prediction and profitability is also studied.

Churn prediction is a binary classification problem; the authors of [6] specified that from the studies it has been observed that there is no proper means of measuring the certainty of the classifier that has been employed for churn prediction. It has also been observed that the accuracy of the classifiers differs for different zones of the dataset. Experimental observations have been made by the authors and it reveals that the distance factor plays a key role in deciding the certainty of the classifier.

[7] Designed a new approach called as logit leaf model which combines the concepts of logistic regression and decision tree model. The proposed approach considers various segments of the dataset rather than considering the entire dataset. This model proves to give better result than the traditional logistic regression and random forest models.

The authors of [8] have studied various machine learning models such as support vector machine, K- Nearest neighbour, Random forest and ADA Boost for their performance in churn prediction. The data considered for designing the model is customer transaction and the authors conclude that the available machine learning model is enough for churn prediction. But the accuracy obtained is less.

Deep artificial neural networks are used in [9] for predicting customer churn in telecommunication industry. 30 features were considered and experimentation are done with various values required for artificial neural networks and produced a result of 85% accuracy.

In [10] customer churn prediction is done with extreme learning machines, a kind of neural networks. 21 features are used in the model but the accuracy obtained is very low compared to the other state of art models.

In order to better utilize the features of the times series data, Long short term memory and convolutional neural networks are used in [11]
.XGBoost is also employed by the authors to extract new features. The accuracy obtained with this approach is 87%.
The customer behaviours are captured as dynamic data and the geographic details are considered as the static data in [12] and they are used for training and testing the churn prediction model designed with feature embedded CNN. These dynamically extracted features produces better result when compared with the conventional models such as LLR, SVM , RF and NN that employs hand crafted features.
The authors of [13] have employed various machine learning models for predicting the customer churn. The model in addition to predicting the churn also identifies the reason behind the customer churn.
[14] Generates time series data from telecom industry with customer conversation and employs similarity forest model to predict customer churn. It provides better result than the existing models.
[15] combines classification and clustering in order to provide a better churn prediction model and the suggestive actions to be taken thereof . Various classification models are analysed and it is found that random forest out performs the other models with 88.63%. The classification process is continued with clustering which gives deeper insights for the organization to suitable and necessary actions.
In addition to the various machine learning models there were also earlier models that uses personal influences for churn prediction [16] and one that combines particle swarm optimization algorithm [17] with the traditional classifiers for churn prediction.

III System Model

Experiment is conducted with a dataset consisting of the details of the bank customers. The dataset contains details of 10000 customers classified into customers who exit represented with 1 in the target variable and 0 if not exited.
The dataset contains 2037 customers who leaves the bank and the remaining continues in the bank. The fields of individual customers include geography, credit score, Gender, Age, Tenure, Balance, Number of products, values specifying whether the customer has credit card, value specifying whether the customer is an active member or not and at last the customer's estimated salary.
There is also a field that represents the target variable. The following figure 1 represents the steps involved. Before applying the models, preprocessing is done in order to convert the categorical data into numerical data.
.

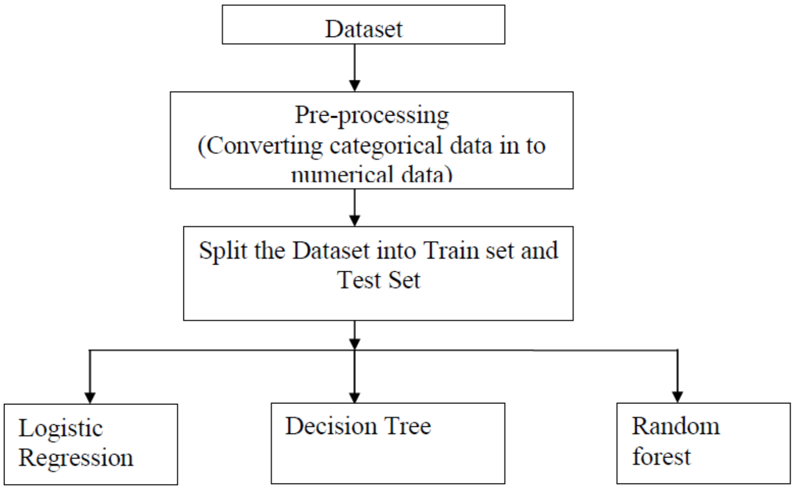


Figure 1 steps in machine learning experiments

Converting categorical data into numerical data is done with label encoders and one hot encoders available in the python. After this conversion is done, the dataset is split into train set and test set and given as input to the various classifiers. The ration followed is 80:20. 80 percentage of the dataset is used as training data and the remaining 20 percentage of the data is used as test data. The results obtained are studied comparatively. The classifiers that are tested are

- Logistic Regression
- Decision Tree
- Random forest.

Along with these baseline models, a model based on artificial neural networks is also proposed. The layers in the proposed artificial neural network are as follows. It contains an input layer, two hidden layers and an output layer. The input layer contains 11units and both the hidden layers contain 6 units. The activation function used in the hidden layers is rectifier linear unit and the activation layer used in the output layer is sigmoid function.
Since we would like to calculate the probability of leaving for individual customer the sigmoid curve is used.
The sigmoid curve is shown in the following figure 2

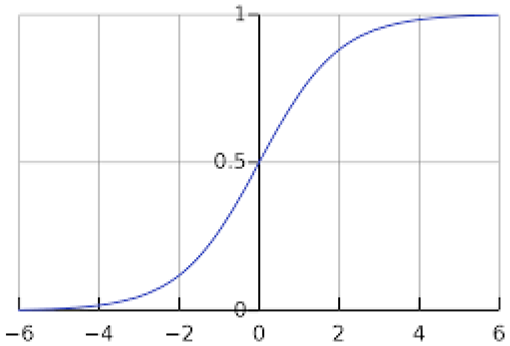


Figure 2 Sigmoid curve

The sigmoid equation is $(x) = \frac{1}{1+e^{-x}}$

The value obtained through this sigmoid function is not 0 or 1 like other activation functions, it lies between 0 and 1. The value specifies the percentage of probability from which the customers can be ranked.
The parameters that are used for comparing the above said classifiers are given below.

(1)True Positive Rate (TPR) = $\frac{TP}{TP+FP}$

Where, TP =True Positive, FP =False Positive

False Positive Rate (FPR) = $\frac{FP}{FP+TN}$

(2)
Where TN=True Negative
Precision = $\frac{TP}{TP+FN}$

(3)
Where FN is False Negative

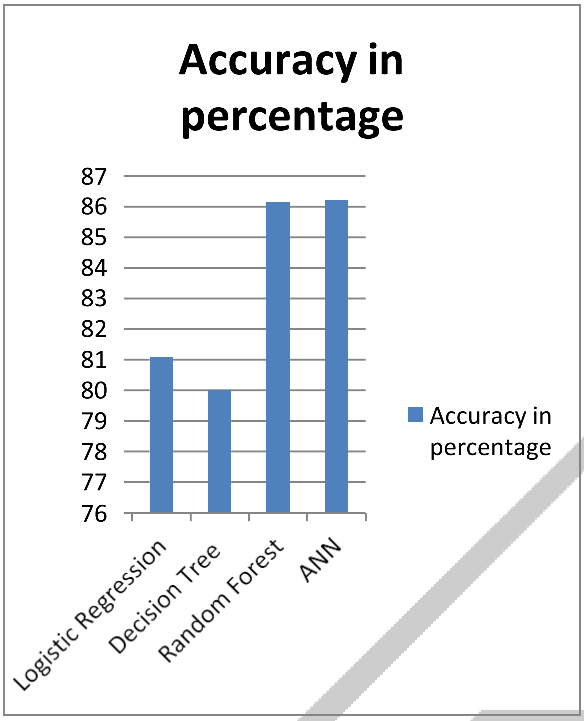
(4) Recall = $\frac{TP}{TP+FP}$
Accuracy= (TP+TN)/ (TN+FP+FN+TP)

The results obtained from the different models and the comparisons of the results are provided in the next section.

IV Results

The experimentation is done with a system that has i3 processor and 4 GB RAM. The models are implemented with sci-kit library.

The following figure 3 represents the accuracy obtained with the individual models.



The following table 1 represents the precision, recall, F1-Score of the three models.

-	cl ass	Logistic regre ssion	Decision Tree	Random fore st	A N N
Precision	0	0.83	0.89	0.88	0.8 9
	1	0.58	0.51	0.73	0.6 9
Recall	0	0.96	0.86	0.95	0.9 4
	1	0.24	0.57	0.50	0.5 6
F1	0	0.89	0.87	0.92	0.9 1
	1	0.34	0.54	0.59	0.6 1

The following figure 4 represents the precision of the three models for both the classes churn and non- churn.

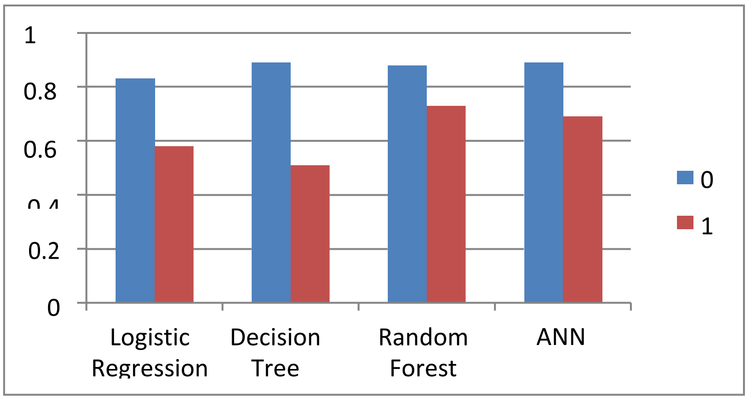


Figure 4 Precision values for both the classes (LR, DT,RF and ANN)

Figures 5 and 6 represent the recall and F1-Score of all the models.

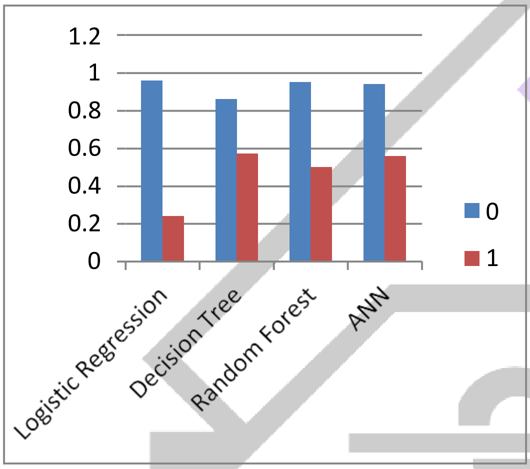


Figure 5 Recall values for both the classes (LR, DT , RF and ANN)

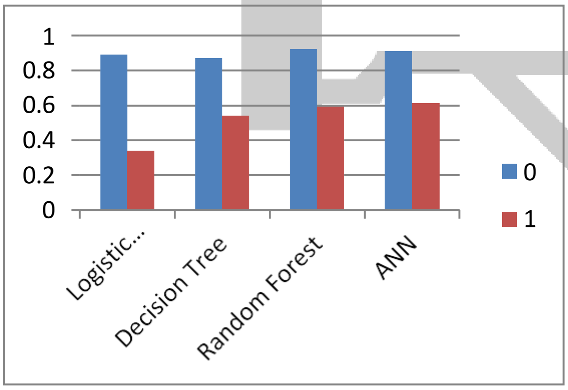


Figure 6 F1-Score values for both the classes (LR, DT , RF and ANN)

From the results it is observed that artificial neural network performs more or less equal to the random forest model and both performs better than the other two models in terms of accuracy, precision, recall and F1 score.

IV Conclusion

A Deep artificial neural network is proposed to predict the customer churn in banking domain. In addition to viewing it as a classification problem, ranking of customers based on the probability of them leaving the organization is also made. Although studies have shown that, deep learning models are also employed for churn prediction, no model has a mechanism to rank the customers in an order so that the proper precautionary mechanisms can be taken by the organization. The model is compared with the three baseline models and found to perform well.

References

[1] V. L. Miguéis, D.vanden Poel,
A. S. Camanho, and J. F. e Cunha,2012,``Modeling partial customer churn: On the value of _rst product-category purchase sequences," *Expert Syst. Appl.*, vol. 12, no. 12, pp. 11250-11256.

[2] A.Idris and A.Khan,2012,``Customer churn prediction for telecommunication: Employing various various features selection techniques and tree based ensemble classi_ers," in *Proc.15th Int. Multitopic Conf.*, pp. 23_27.

[3] W. Verbeke, D. Martens, C. Mues, and B.Baesens,2011, ``Building comprehensible customer churn prediction models with advanced rule induction techniques," *Expert Syst. Appl.*, vol. 38, no. 3, pp. 2354-2364.

[4] Farid Shirazi, Mahbobeh Mohammadi,2019,"A big data analytics model for customer churn prediction in the retiree segment",International Journal of Information Management,Volume 48,Pages 238-253.

[5] A.DeCaigny ,K.Coussement, K.W. De Bock et al.,(2019) Incorporating textual information in customer churn prediction models based on a convolutional neural network. International Journal of Forecasting .

[6] Adnan Amin,Feras Al-Obeidat, Babar Shah, Awais Adnan, Jonathan Loo, Sajid Anwar,2019,"Customer churn prediction in telecommunication industry using data certainty",Journal of Business Research,Volume 94,pp 290-301.

[7] Arno De Caigny, Kristof Coussement, Koen W. De Bock,A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees,European Journal of Operational Research,Volume 269, Issue 2,2018,pp760-772.

[8] Satu Elisa Schaeffer, Sara Veronica Rodriguez Sanchez,2020,"Forecasting client retention — A machine-learning approach",Journal of Retailing and Consumer Services,Volume 52,2020.

[9] Nigam, B., & Dugar, H.2019, Effectual Predicting Telecom Customer Churn using Deep Neural Network.International Journal of Engineering and Advanced Technology.Volume 8 issue 5, pp121-127

[10] Mohanty R., Naga Ratna Sree C. (2018) Churn and Non-churn of Customers in Banking Sector Using Extreme Learning Machine. In: Bhateja V., Tavares J., Rani B.,Computational Intelligence and Informatics. Advances in Intelligent Systems and Computing, vol 712.Springer, Singapore

[11] ZHOU, Jie & YAN, Jian-feng & YANG, Lu & WANG, Meng & XIA, Peng. (2019). Customer Churn Prediction Model Based on LSTM and CNN in Music Streaming. DEStech Transactions on Engineering and Technology Research. pp 254-261

[12] Chongren Wang , Dongmei Han , Weiguo Fan and Qigang Liu,2019
," Customer Churn Prediction with Feature Embedded Convolutional Neural Network: An Empirical Study in the Internet Funds Industry", International Journal of Computational Intelligence and Applications, Vol. 18, No. 1.

[13] Beschi Raja ,Pamina Jeyakumar(2019), "An effective classifier for predicting churn in Telecommunication" , Journal of Advanced Research in Dynamical and Control Systems, Vol. 11, 01-
Special Issue, 2019,pp 221-229

[14] María Óskarsdóttir, Tine Van Calster, Bart Baesens, Wilfried Lemahieu, Jan Vanthienen,2018,"Time series for early churn detection: Using similarity based classification for dynamic networks",Expert Systems with Applications,Volume 106,
2018,Pages 55-65.

[15] I. Ullah, B. Raza, A. K. Malik, M. Imran, S. U. Islam and S. W. Kim, 2019, "A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector," in *IEEE Access*, vol. 7, pp. 60134-60149.

[16] XiaohangZhang,JiZhu,Shuhua Xu,YanWan,(2012),Predicting customer churn through interpersonal influence, Knowledge-Based Systems,Volume 28, 2012,pp 97-104.

[17] Adnan Idris, Muhammad Rizwan,Asifullah Khan,2012,"Churn prediction in telecom using Random Forest and PSO based data balancing in combination with various feature selection strategies",Computers & Electrical Engineering,
Volume 38, Issue 6,2012,pp 1808-1819.