Customer Churn Prediction

Final Project Report

Madiraju Harsha Nandan

Project Objective

In today's competitive world in telecom industry, customer churn is one of the main concerns. In this domain, churn is described about the customers who ends utilizing the services and change to the other network. This results in possible loss of income for the company. Furthermore, it has become a compelling issue to maintain customers therefore, it is essential to recognize the customers who may potentially leave the network in a short period. This process is described as churn prediction.

The data indicates the customers who churned their service. Based on past data, we need to build a classifying model which can predict whether a customer will cancel their service in the future or not. The pipeline followed to implement the model is as follows.

- Importing the dataset in python
- Understanding the structure of dataset and modifying the data into format
- Graphical exploration
- Descriptive statistics
- Insights from the dataset
- Check if the assumptions are met
- Creating models using deep learning with the help of tensor flow
- Validating the models with various performance measures (Confusion matrix, ROC/AUC)
- Actionable insights and Recommendations

1. Background and Related Work

Deep learning is a sub-domain in artificial neural networks (ANNs), which consists of many hidden layers. Every layer in deep learning algorithms will be trained on all data features came from previous layer of the network. Along these lines, an element order of expanding deliberation can be made, which is great at finding designs inside a dataset. In this neuron are interconnected with large numbers of neurons and they are managed in the form of layers. As we term layers there are comprises of many layers that will frame the total brain organization. There have been uses of beat displaying in retail space, just the examinations that are of the area and with RFM center as well as profound learning applications in stir setting will be checked. Churn studies mostly based on contractual setting sectors such astelecommunication, banking, insurance, etc. where a customer must sign a contract to get the service from the provider. Subsequently, a client who ends the agreement marked as a churner by the organization though the client who keeps on getting administration acknowledged as a non-churner.

2. Methodology

In this study, a telco customer churn dataset was used. The data set has 7,044 records and 21 variables and includes information about:

- Customers who left within the last month, the column is called Churn
- Administrations that every client has pursued, telephone, numerous lines, web, online security, online reinforcement, gadget insurance, technical support, and streaming TV and films
- Client account data, how long they've been a client, contract, installment strategy, paperless charging, month to month charges, and absolute charges
- Demographic info about customers, gender, age range, and if they have partners and dependents proposed

A neural network model was implemented using the data with the same classes of churn or non-churn. The 80% of data is used in the training set (5625) and remaining 20% of data is used in the test set (1407). So, there is an 80%-20% apportionment and all the prediction models used this proportion for the validation.

3. Performance Metrics

Accuracy, precision, recall, and area under the curve (AUC) were used to evaluate the performance of the model. For extracting these metrics, a confusion matrix was used. A confusion matrix is a simple table to that explains ratios of the performance of a classification/prediction model on which the values are known. Additionally, it assists with picturing the presentation of the model utilizing Receiver Operations Curve. Confusion Matrix consists of True Positive (TP), False Positive (FP), False Negative (FN), True Negative (TN).

Important Measures of Confusion Matrix: Sensitivity, Specificity, Accuracy

Sensitivity: Also called as True certain rate or Recall. This is extent of real sure cases which are accurately recognized.TP/(TP+FN)

Specificity: Also called as True Negative rate or False Positive rate. This is proportion of negatives that were correctly identified. TN/(TN+FP)

Accuracy: 1-error rate. This is the number of right expectations are done in the two classes. Errorrate: FP+FN/(TP+TN+FP+FN)

From the data,the customers who got churned are marked as ‘1’, customers who didn’t churn are marked as ‘0’.

Below is the accuracy for ANN Model.

Model	Accuracy
ANN	0.811

Artificial Neural Network with keras:

The dataset is relatively huge, on train dataset achieved the accuracy of 81.1% and on test dataset the accuracy is 80.1%.

The precision on train dataset is 70.21% and on test dataset it is 71.32%.

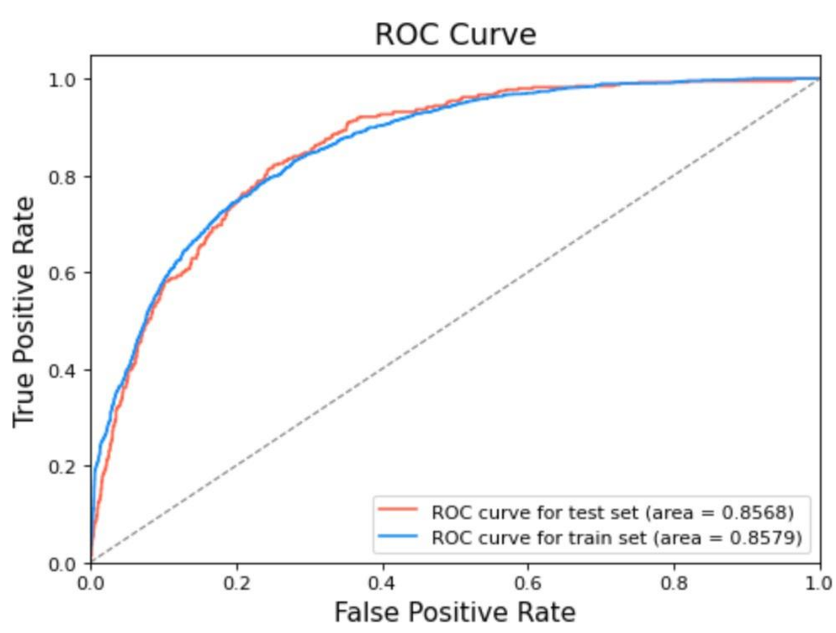
The recall on train dataset is 49.19% and on test dataset it is 48.09%. This model did good predictions.

4. Interpretation of other Model Performance Measures(AUC):

AUC-ROC:

ROC and AUC stand for Receiver Operating Characteristics and Area under the curve.

AUC-ROC Curve:



Area under curve for test data is 85.68. ROC is plotted axis are below:X axis: False positive rate/ Specificity=FP/(FP+TN)
Y Axis: True Positive rate/ Sensitivity/Recall=TP/(TP+FN)

Usually, the area under test data curve is relatively less than train data curve. TPR will increaseand becomes stagnant eventually. Higher the area under the curve, better is the performanceof the model.

5. Conclusion

So far, customer churn prediction studies were conducted to reveal which customers are probably to leave the network. Several models and prediction techniques have been used. The deep learning methods, as one of the recent developments within the scope of artificial neural networks and are frequently used in image processing and image definition, can also be utilized in churn studies. In this review, a profound learning model in client stir expectation with crate examination highlights was introduced. Models' exhibitions were estimated with the disarray grid yields including exactness, review, accuracy, and AUC. According to these performance metrics, the DL model with the inclusion of promotional data revealed higher accuracy of churn prediction.