IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 22, Issue 3, Ser. I (May - June 2020), PP 54-59
www.iosrjournals.org

Predicting Customer Churn in Telecommunication Industry Using Convolutional Neural Network Model

Sunday A. AMATARE[1], Adebola K. OJO[2]*
[1,2]*Department of Computer Science, University of Ibadan, Nigeria.*
[*]*Corresponding author*

Abstract: *In this study a Convolutional Neural Network (CNN) model was proposed for the prediction of customer churn in a telecommunication industry. Many supervised machine learning models have been built and used for predicting customer churn in past researches. However, in thebuilding of these models, there is need for human intervention to carry out attributeselectionwhich is very tedious, time-consuming, tailored to specific datasets and often result to attribute selection problems.*

This study proposed a convolutional neural network model for predicting customer churning behavior and to also get rid of human attribute selectionand its problems. Two datasets were created from the fourteen thousand data instances that were gotten from one of the major cellular companies operating in Nigeria. Python programming language via the anaconda distribution was used for the development and implementation of our model. Jupyter notebook was our IDE choice. In other to achieve a like-for-like comparison, three other models were developed, which were two Multi-layer Perceptron (MLP) models and one other CNN model. The accuracy rates for the MLP models; MLP1 and MLP2, are 80% and 81% respectively while the CNN models, CNN1 and CNN2, are 81% and 89% respectively.

Keywords*:Convolutional Neural Network, Customer Churn, Attribute selection, Multi-layer Perceptron.*

Date of Submission:02-05-2020 Date of Acceptance: 16-05-2020

I. Introduction

Some decades ago, there has been proliferation of data due to the advancements of technology. However, data doesn't speak for itself; it has to be spelled out. A wide range of methods and techniques have been used for the selection, preparation and processing of data so as to uncover hidden and interesting information from it. The process is known as Data Mining[1].

Customer churn is when a user or a subscriber stops using a company's product or service. In telecommunication paradigm, churn means subscribers' movement from one service provider to another. It results to possible loss of business. For this reason, telecom service providers are engaging all manner of strategies to tie down their customers. Prior to that, it is imperative for them to know the customers that are likely to opt out of the company because opting out means loss of business for the company[2].

In 2013, MTN, one of the leading cellular companies operating in Nigeria, launched the porting service in Nigeria as a requirement by the Nigerian Communications Commission (NCC), and several other telecommunication companies followed suit. The porting service offers subscribers the opportunity to use the services of another provider and still maintain their cell number. It is a game-changer because it provides a range of options for users and promotes effective competition by giving subscribers the freedom to move from one company to another and still maintain their cell number[3]. Customer is the source of profit for telecommunication companies. The loss of customers is the loss of revenue/business. The cost of acquiring a new customer is 5 to 6 times higher than the cost of retaining an existing customer [4]. Because of the financial implication, telecommunication companies have moved their attention away from winning new customers to retaining loyal customers. In other for them to be successful and competitive in the saturated market, there is need for them to be able to forecast potential churn and take timely retentive efforts. CNN is a specific type of artificial neural network that uses perceptrons, a machine learning unit algorithm, for supervised learning, to analyze data. CNN have been used in variety of areas, including image and pattern recognition, natural language processing, and video analysis.

In this paper, a predictive analysis of churning behavior was done in a telecommunication industry based on the dataset that was extracted in one of the major telecommunication companies in Nigeria. The model was built using the architecture of convolutional neural network; customers who are likely to quit their service or not were predicted, we used standard classification performance metrics to measure the robustness of the model. Also, two datasets were created, and four models were developed using our datasets to achieve a like-for-like comparison. Our paper was arranged in this order. Section 2 talks about our study related works. Section

3 explains our methodology. Section 4 presents our results and findings. Section 5 talks about conclusion and future work.

Related Works

Because winning new customers is much costlier than keeping existing customers, predicting customer who may likely opt out of a company has become very crucial for all telecommunication industries in planning and making decisions. A model was built to predict customer churn in the telecommunication sector using Rough Set Approach. 3,333 data instances were used and Weka was employed for the selection of relevant features. Experiment was carried out to assess the effectiveness of the four types of algorithm used. The result showed that genetic algorithm gave a better result than the three other algorithms. The technique was able to correctly predict customers who may likely opt out of a service and gave information for retentive efforts.[5],developed a model for predicting customer churning behavior in a telecommunication in a telecommunication industry using a clustering algorithm. The study used C5.0 and back propagation neural network classification algorithms to predict churn. Self-Organizing Map technique was employed to group subscribers into different clusters according to their service behaviors. [6] worked on a multilayer perceptron model for the prediction of customer churn in a financial institution. The dataset used has 50000 data instances with 42 features. Dropout and L2 Regularization techniques were used to ensure that the model learns correctly. Python and Neuro Solution Infinity Software were used to implement the model and they produced outstanding results of 97.53% and 97.36% respectively.[7] worked on a model for predicting subscribers churning in cellular network services using Neural Network based approach. The dataset contained 2427 data instances with 20 features that were extracted from University of California, Irvine. The implementation of the neural network model was done on Clementine data mining software package from SPSS. The model gave an outstanding result with an accuracy of 92.35%.In the work of [8], customer preference was predicted based on previous customer behavior, mobility characteristics, and their social network activities. The dataset used contained 120,825 restaurant check-ins with 5 features. Artificial Neutral Network (ANN) and Support Vector Machine (SVM) were built and comparison was carried out to check their performance. The ANN model has an accuracy of 93.13% against 54.00% SVM accuracy.[9] developed a hybrid model for predicting churn in mobile telecommunication. Logistics Regression and voted Perceptron were used for classification and clustering. The experiment was carried out on Weka, a well-known machine learning tool that contains 2000 instances and 23 variables from an Asia mobile operator. The result of the model after evaluation gave a superior accuracy when compared with one model.[10]proposed a model for predicting customers that are planning to abandon the service of a provider to port over to other providers using Multilayer Perceptron architecture of Artificial Neural Network. Dataset containing 20468 customers' record with 26 features was gotten from all cellular companies in Pakistan. IBM SPSS Statistics was employed in building the architecture. Back propagation technique was used in training the model. The proposed model has an accuracy rate of 79%. In[11], Artificial Neural Networks (ANN) and Regression Analysis models were considered to determine which of the models performed better. Prediction was done using one hidden layer and three processing elements in the ANN model. The prediction was done using regression analysis. The parameters of regression model were estimated using Least Square method. To determine the better prediction, mean square errors (MSE) attached to ANN and regression models were used. Seven real series were fitted and predicted with in both models. It was found out that the mean square error attached to ANN model was smaller than regression model which made ANN a better model in prediction. In the work of [12],a model was developed for predicting customers that may likely opt out of a service or not using multilayer architecture of Artificial Neural Network. Two methods were used by changing the number of epochs and neurons in the hidden layer. Dataset of 50,000 data points with 11 attributes was collected from one the telecom giants in Jordan. Some standard classification performance metrics were employed to assess the models.

II. Material And Methods

The methodology of our proposed convolutional neural network model was presented in this section. This section includes; data collection, data analysis, data preparation, tools used, and implementation.

Data Collection

Fourteen thousand data instances with 15 attributes were extracted from one of the leading cellular companies operating in Nigeria. Table 1 presents the dataset description.

Table 1: Description of Data Used

Attributes	Description
Customer_Id	Unique Identification Number for customers.
Network_age	The period of time customer has been using the service.
Customer_tenure_in_months	The number of months that a customer has subscribed for.

Total_Spend_in_Months_1_and_2_of_2017	The total amount of money in Naira spent by a customer in month 1 and 2 in 2017
Total_SMS_Spend	The total amount of money in Naira a customer spent on SMS.
Total _Data_Spend	The total amount of money in Naira a customer spent on data.
Total_Data_Consumption	The total volume of data a customer used.
Total_Unique_Calls	The total number of unique calls a customer made.
Total_Onnet_Spend	Total amount of money spent by a customer on calls or messages on the same network provider.
Total_Offnet_Spend	Total amount of money spent by a customer on calls or messages on different network provider.
Total_Call_center_complaint_calls	Total number of calls a customer made to the network Call center for complaints.
Network_type_subscription_in_Month_1	The generation of network (2G or 3G) a Customer made subscription in Month 1.
Network_type_subscription_in_Month_2	The generation of network (2G or 3G) a customer made subscription in Month 2.
Most_Loved_Competitor_network_in_Month_1	Customer alternate service provider in Month 1.
Most_Loved_Competitor_network_in_Month_2	Customer alternate service provider in Month 2.

Data Preparation

After analysis of the dataset, we discovered that there were missing values in some cells. We were able to create two datasets out of it. In the dataset, categorical data were turned to numerical data. The dataset was normalized for faster computation. The dataset was partitioned into two parts, 80% and 20%, for training and testing the model respectively. For data 1, we removed the column (most loved competitor network in month 1) because it had plenty of missing values. For data 2, we removed rows (data points) with missing values. We used data 1 and 2 to build to CNN models; we also used data 1 and 2 to build two MLP models so as to achieve a like-for-like comparison. Below is a sample of some data points in the customer churn dataset.

Tools and Libraries

The following tools were used in the analysis of the data, extraction of features, development and evaluation of the machine learning model: Python programming language via the anaconda distribution of the programming language was used for all development and implementation. Jupyter notebook is the IDE of choice. The libraries and packages used are; Pandas, Pandas Profiling, Numpy, Keras, Sci-kit learn, Seaborn, and Matplotlib.

Implementation

Python programming language was used to develop our convolutional neural network model. Our convolutional neural network architecture is made up of the following:

Conv1D (2)→Dropout (0.2)→MaxPool (2)→ Flatten→ Dense (10)→ Dense (1)

- Convolution layer with kernel size of 2
- Dropout layer of 20%
- MaxPool layer with pool size of 2
- A flatten layer
- Fully connected layer with 10 nodes
- Fully connected layer with 1 node which is the output.

Steps for Proposed CNN Model

- Extraction of customers' data from the database of a telecommunication company.
- Data preprocessing for data cleaning and data transformation.
- Normalization for standardization and faster computation.
- Data partitioning (train and test set).
- Build our CNN model
- De-normalization of the output for interpretation.
- Predict churn or not churn.
- Evaluate the model.

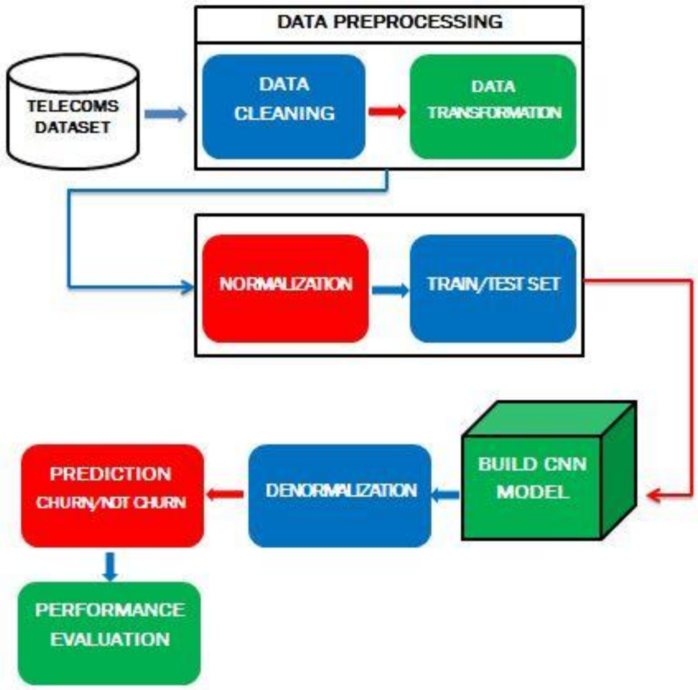

Figure 1: Proposed CNN model

III. Results and Findings

Our convolutional neural network model result that was implemented in Python programming language was shown in Figure 2. Using our training dataset, it has an accuracy rate of 88.22%.

```
1  model2_2.fit(xtrain2, yLabel_train2, batch_size=10, epochs=100, verb
 - 0s - loss: 0.4379 - acc: 0.8827 - val_loss: 0.5484 - val_acc: 0.8914
Epoch 93/100
 - 0s - loss: 0.4342 - acc: 0.8745 - val_loss: 0.5832 - val_acc: 0.8851
Epoch 94/100
 - 0s - loss: 0.4373 - acc: 0.8991 - val_loss: 0.5480 - val_acc: 0.8637
Epoch 95/100
 - 0s - loss: 0.4380 - acc: 0.8970 - val_loss: 0.5430 - val_acc: 0.8737
Epoch 96/100
 - 0s - loss: 0.4306 - acc: 0.8934 - val_loss: 0.5461 - val_acc: 0.8822

Epoch 97/100
 - 0s - loss: 0.4312 - acc: 0.8991 - val_loss: 0.5718 - val_acc: 0.8679
Epoch 98/100
 - 0s - loss: 0.4345 - acc: 0.8724 - val_loss: 0.5367 - val_acc: 0.8865
Epoch 99/100
 - 0s - loss: 0.4373 - acc: 0.8949 - val_loss: 0.5465 - val_acc: 0.8750
Epoch 100/100
 - 0s - loss: 0.4305 - acc: 0.8845 - val_loss: 0.5422 - val_acc: 0.8822

<keras.callbacks.History at 0x248025fd588>
```

Figure 2: Learning Process

Performance Evaluation:

We employed the use of some standard classification performance metrics like accuracy, precision, recall, f1-score and confusion matrix to check the effectiveness and robustness of the model.

```
### Predict and evaluate using CNN on data2
cnn_data2_predict = cnn_data2.predict(np.expand_dims(data2_x,  axis=2))
cnn_data2_pred = np.around(cnn_data2_predict)
conf4 = confusion_matrix(data2_y, cnn_data2_pred)
class4 = classification_report(data2_y, cnn_data2_pred)
print(conf4)
print(class4)
```

```
[[511  56]
 [ 68 534]]
              precision    recall  f1-score   support

           0       0.88      0.90      0.89       567
           1       0.91      0.89      0.90       602

    accuracy                           0.89      1169
   macro avg       0.89      0.89      0.89      1169
weighted avg       0.89      0.89      0.89      1169
```

Figure 3: CNN – Confusion matrix, Accuracy, Precision, Recall and F1-Score.

Receiver Operation Characteristic (ROC) curve was used. The ROC curve gave us an excellent result. The value of our ROC curve was 0.89.

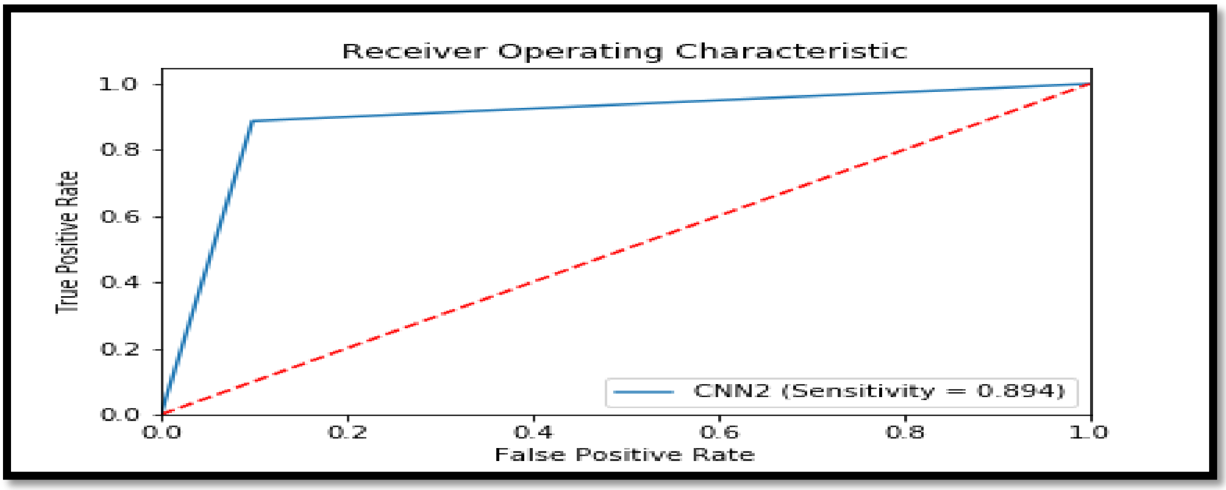

Figure 4: CNN ROC curve.

Models Evaluation:

In order to achieve a like-for-like comparison, we built additional 3 models: two Multi-layer Perceptron (MLP) models and a Convolutional Neural Network (CNN) model, using the two datasets (dataset 1 and 2) created from our original dataset that was extracted from one of the leading cellular companies operating in Nigeria. Standard performance metrics such as accuracy, precision, recall, f1-score, and confusion matrix were performed on the models to know their effectiveness and robustness. The models achieved the following results:

Table 2: Models results

	MLP-1	MLP-2	CNN-1	CNN-2
Accuracy	0.80	0.81	0.81	0.89
Precision	0.80	0.81	0.81	0.89
Recall	0.80	0.81	0.81	0.89
F1-score	0.80	0.81	0.81	0.89
Confusion Matrix	[[462 128] [113 498]]	[[453 114] [113 489]]	[[471 119] [108 503]]	[[511 56] [68 534]]

Customer Churn Prediction Results

The resultant model after training produced the following prediction results placed side-by-side for the same data point.

```
import keras
from keras.models import load_model

data = [75.07,535.63,1.75,7.5,1919857.372,131,8820,23657,2,'Other','O

#load model from file
mlp_data1 = load_model('churn-net-d1-v01-780acc.h5')
mlp_data2 = load_model('churn-net-d2-v01-735acc.h5')
cnn_data1 = load_model('churn-cnn-d1-net-v01-751acc.h5')
cnn_data2 = load_model('churn-cnn-d2-net-v01-722acc.h5')

mlp_data1_predict = mlp_data1.predict(data_for_model1)
mlp_data2_predict = mlp_data2.predict(data_for_model2)
cnn_data1_predict = cnn_data1.predict(np.expand_dims(data_for_model1,
cnn_data2_predict = cnn_data2.predict(np.expand_dims(data_for_model2,

print('MLP1: ',mlp_data1_predict[0][0]*100,'% probability of churn')
print('MLP2: ',mlp_data2_predict[0][0]*100,'% probability of churn')
print('CNN1: ',cnn_data1_predict[0][0]*100,'% probability of churn')
print('CNN2: ',cnn_data2_predict[0][0]*100,'% probability of churn')
MLP1:  18.656612932682037 % probability of churn
MLP2:  3.525986894965172 % probability of churn
CNN1:  10.529936105012894 % probability of churn
CNN2:  5.877472460269928 % probability of churn
```

Figure 5: Models results on same sample point.

IV. Conclusion

Customer churn prediction has become more important task because of the competition and freedom of the cellular market. Early forecast of loyal customers who may quit the service of a company can bring about preventive actions which can help in retaining them. Consequently, it is imperative for telecom companies' owner to build an accurate and precise churn model.

We built a Convolutional Neural Network model for predicting customer churn in telecommunication industry. It showed that CNN, which is mostly used for image classification, can also be used for churn prediction. Our model gave excellent accuracy of 89%. Also, excellent results were gotten from the three other models (two MLPs & one CNN) developed to achieve a like-for-like comparison. We employed the use of some standard classification performance metrics like Accuracy, Precision, Recall, F1-score and confusion matrix on our datasets. For future work, ensemble techniques could be used for getting higher accuracy. Also, other neural network models should be used for customer churn prediction.

References

[1]. A., Saran Kumar; D., Chandrakala, "A Survey on Customer Churn Prediction Using Machine Learning Techniques," International Journal of Computer Applications, Vols. 154 - No. 10, pp. 0975 - 8887, November 2016.

[2]. V., Umayaparvathi, K., Iyakutti;, "Attribute Selection and Customer Churn Prediction - Deep Learning Approach," 2016.

[3]. A. Boateng and O. O. Owusu, "Mobile Number Portabilty: On the Switching Trends among Subscribers within Telecommunication Industry in a Ghanian City," 2013.

[4]. C. B. Bhattacharya, "When Customers are Members: Customer Retention in Paid Membership Contexts," Journal of the Academy of Marketing Science, 1998.

[5]. Zhang, Yongbin; Liang, Ronghue; Zheng, Yeli Li Yanying; Berry, Michael, "Behavior Based Telecommunication Churn Prediction with Neural Network Approach," International Symposium on Computer & Society, 2011.

[6]. Amuda, Kamorudeen A., Adeyemo, Adesesan B., "Customer Churn Prediction in Financial Institution using Artificial Neural Network," 2019.

[7]. Sharma, Anuj; Panigrahi, Prabin Kumar, "A Neural Network based Approcah for Predicting Customer Churn in Cellular Network Services," International Journal of Computer Applications, Vols. 27, No. 11, August 2011, 2011.

[8]. B. Zheng , K. Thompson, S. S. Lam and S. W. Yoon, "Customers' Behaviour Prediction Using Artificial Neural Network," in Industrial and Systems EngineerimgReszearch Conference, 2013.

[9]. Olle, Georges D. Olle; Cai, Shuqain, "A Hybrid Churn Prediction Model in Mobile Telecommunications Industry," International Journal of e-Education, e-Business, e-Management and e-Learning., Vols. 4, No. 1, February 2014, 2014.

[10]. Y. Khan, S. Shafiq, A. Naeem, S. Ahmed, N. Safwan and S. Hussein, "Customer Churn Prediction using Artificial Neural Networks (ANN) in Telecom Industry," International Journal of Advanced Computer Science and Applications (0975 - 8887), Vols. 10, , no. No. 9, 2019, 2016.

[11]. Ojo A.K. and Adeyemo A.B., "A Comparison of the Predictive Capabilities of Artificial Neural Networks and Regression Models for Knowledge Discovery," Information Systems, Development Informatics and Business Management, vol. 4, no. 2, pp. 27-32, 2013.

[12]. Adwan, Omar; Faris, Hossain; Jaradat, Khalid; Harfonshi, Osama, "Predicting Customer Churn in Telecom Industry using Multilayer Perceptron Neural Networks: Modelling and Analysis," Life Science Journal, pp. 11(3):75-81, January 2014.