

Churn Prediction on Telecommunication Services Using Artificial Neural Network

Sandeep Kaur

Assistant Professor

School of Computer Science and Engineering

Lovely Professional University, Punjab

sandeep.23614@lpu.co.in

Danish Farooq

School of Computer Science and Engineering

Lovely Professional University, Punjab

farooqdanish101@gmail.com

Mustahseen Shafi

School of Computer Science and Engineering

Lovely Professional University, Punjab

mustahseen66@gmail.com

Gurpreet Singh

School of Computer Science and Engineering

Lovely Professional University, Punjab

gurpreet26756@gmail.com

Abstract— During the past few decades, mobile communication has been the dominant medium of communication. In developed countries, with increasing number of telecom operators and rival competitors, churn prediction has become a major Business Intelligence(BI) application that aims at identifying customer retention rate. The mobile sector is affected immensely due loss of revenue when customer is likely to cancel a subscription and switch to competitor. It is important for business analysts to know the reasons for churn customers and identify behavioural patterns associated with it. This paper presents a model based on artificial neural network which tries to find patterns using historical data for prediction of customer's churn. This work reveals the factors behind churning of churn customers so that telecom companies can work upon customer retention.

Keywords— Churn Analysis, Customer Retention, Marketing, Churn.

1.Introduction

Churn prediction uses a standardized mechanism for identifying and analyzing customer patterns and trends. The patterns are used by the telecom companies to improve their customer retention and thereby improving their profits [5]. The identification of churns and non churners is tedious task and needs to be carried out in

time as a lot depends upon this. This job is considered tedious because of following reasons:

(i) The Customer data is huge.

(ii) Available data is not complete moreover inconsistent too thus making the task of formal analysis a tedious one.

Moreover, because of its enormous size, the investigation and analysis of the database that contains customer data takes considerably long amount of time due to the sophistication of these issues. As information science has progressed complex data mining and AI is easily available to the telecom sector [6,7]. The practices when combined with computers can process thousands of statements in lesser amount of time, thus saving a lot of time. Furthermore, the installation and software costs less than training and hiring. Human investigators are prone to errors and lack efficiency whereas computers are less prone to errors and can used to work for long hours.

What telecom industries need today is a comprehensive tool that can be used to identify customer patterns and help them to locate churners and mechanism through which those churners may be converted to non-churners. Current research work aims developing such a system that can be used by telecom companies to easily discover customer likelihood and trends thereby make predictions, discover relationships and possible explanations and

identify possible churners. The use of data mining and AI is important for design and development. Churn prediction is knowing which of the particular customers will leave a particular service [4,10].

1.1 Scope of Study

The study covers the need and effectiveness of Novel approach for Churn analysis. Our model or approach saves precious time in cleaning the data and analysing the churn. Churn analysis can be helpful in customer retention [8]. Moreover, the accuracy can be enhanced by using a greater number of variables the availability of data and AI tools has furthermore helped in increasing the accuracy but still a lot needs to be done.

2. Churn Analysis

We used Machine learning algorithm namely ANN, Artificial Neural networks is nothing but a computational model that emulates biological neural networks. It is considered as a nonlinear statistical data. There exists a complex relationship between output and input [9].

We use ANN algorithm for the prediction of churn rate. The training dataset and testing data is taken as the percentage of 80% and 20% respectively. Basically, this model works on multiple attributes like facts and figures based on the demography of the customers, details of their billing and consumption designs from telecom data set. In comparison with other prediction techniques which are available, we got the outcome from Artificial Neural Networks based technique quite impressive and it predicted the telecom churn with significantly greater accuracy. The results from ANN clearly depicts the churn factors and now then companies will revise their policies and take required steps so that to decline the cause of churn like devising a predetermined marketing strategy.

3. DATA PREPARATION AND FEATURE EXTRACTION

For our Churn analysis we have a used a publicly available dataset of telecom companies. The feature extraction is an important proponent of our analysis. We need to extract features on which churn depends the most so that we can reduce the computational cost. We deduced the following features from our dataset: Customer id, gender, senior citizen, associate, dependents, phone service, multiple lines, net service, device defense, online protection, total charges churn.

3.1 Visualization

In visualization part, we visualized all the results in the form of graphs using various platforms like Python. For the further visualization, we created a GUI using Python. Moreover, the visualisation part can also be done using Tableau.

3.2 Extracting classification rules from Neural Networks

3.2.1 Using sklearn library

The sklearn library is used to provide sundry utility functions and makes use of transformer classes to change the vector features into a form that is more readable for the downstream estimators.

3.2.2 Using Pandas library

Pandas is an open-source library that is used for data analysis, mainly data manipulation. It is very efficient and provides great performance with back-end code written in Python. It is not a stand-alone library but requires 'NumPy' to operate. It easily resolves the issue of missing data and provides an effective way to slice the data.

3.2.3 Using Keras library

It is also an open-source Neural Network library which is facile to use, quick and modular. It handles high-level computation and like Pandas it also uses another library called "backend". It defines layers and sets up multiple input-output models.

3.2.4 Using matplotlib library

It is a plotting library and is used to plot graphs, histograms etc. It contains a collection of functions with each function making some changes to a figure like plotting some lines in a plotting area, creating a plotting area in a figure etc.

3.2.5 Using Seaborn library

It is a Python library which is used for statistical data visualization. It is a graphic library and allows to make charts fancier. It also facilitates some of the basic visualization needs like mapping et al.

4. Challenges with telecom dataset

It is important to explore the structure of the telecom dataset and thereby make it suitable for machine learning algorithms. During our work we found out that most of variables would amount to 0. So we removed all of them. Since we didn't know the features which would predict the churn we had to work with all of the data.

5. Experimental results and observations

5.1 Reading dataset

```
data = pd.read_csv("WA_Fn-UseC_-Telco-Customer-Churn.csv")
```

5.2 Names of all the attributes

```
>names(data)
"Customer id", "gender", "senior citizen", "associate",
"dependents", "phone service", "multiple lines", "net
service", "online protection", "device defense", "total
charges churn".
```

5.3 Using matplotlib

```
fig,ax = plt.subplots(1, 3, figsize=(20,5))
sns.boxplot(x=data["tenure"], orient="h",
color="purple",ax=ax[0])
sns.boxplot(x=data["MonthlyCharges"], orient="h",
color="olive",ax=ax[1])
sns.boxplot(x=data["TotalCharges"], orient="h",
color="green",ax=ax[2])
```

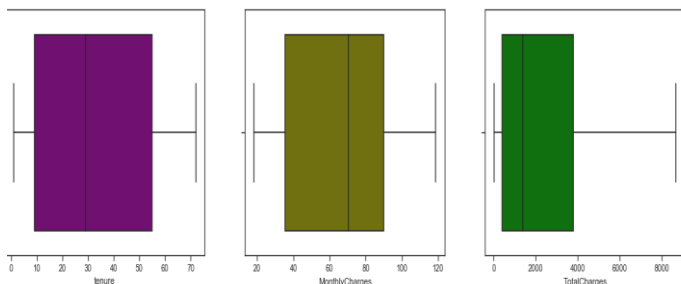


Figure 1: This figure demonstrates the total time the services were availed for and how many monthly and total charges did it account for.

5.4 Bar Graphs

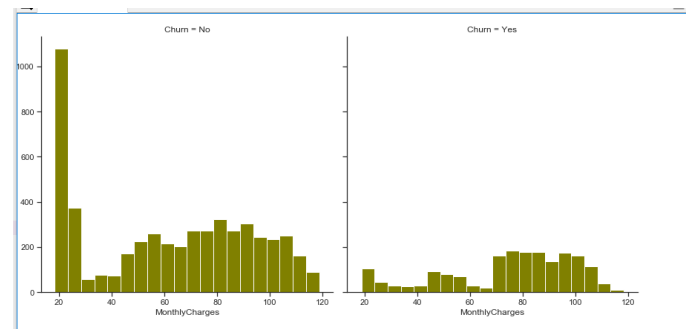


Figure 2: This figure depicts the rate at which the customers are shunning a particular service on a monthly-period.

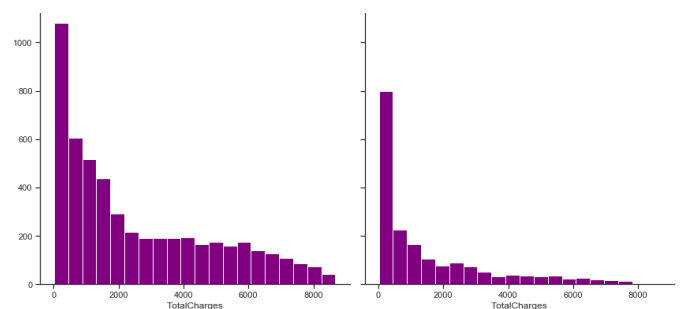


Figure 3: The above figure depicts the difference between the normal usage of services and after churning. It shows the services that the customers are most likely to shun.

6. CONCLUSION

The telecom business during recent times has seen dramatic changes and this industry has grown at an enormous rate. Customer's never ending desires for better services and cost, while their needs and demands are extremely perplexing and difficult to understand. The Churn Factor can be subjected to numerous functions to portrair the various regions or scenarios in cases when the amount of churn rate is maximum. The graphs are figured out taking churn factors as the deciding parameters. These Graphs depicts the dissimilar ways of spotting the amount of churners from the given dataset. From this work, the operators can get to know their customer's behavior more accurately and thus will be able to target policies around customers and focus on their retention. Finally, this work provides guidelines on customer retention for decision-makers of the telecom companies.

7. REFERENCES

- [1] L. Miguel APM. "Measuring the impact of data mining on churn management." Internet Research.
- [2] J. Hadden, et al. "Churn prediction: Does technology matter." International Journal of Intelligent Technology.

- [3] Telecommunication Subscribers ‘Churn Prediction Model Using Machine Learning, Saad Ahmed Qureshi, Ammar Saleem Rehman, Ali Mustafa Qamar, Aatif Kamal.
- [4] L. Breiman, J. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. Wadsworth and Brooks, Monterey, CA, 1984.
- [5] V. Lazarov and M. Capota. Churn prediction, Technische Universit at M unchen. *Eighth ACM SIGKDD International Conference*, 2007.
- [6] Manzano-Machob, ‘‘The architecture of a churn prediction system based on stream mining,’’ in Proc. Artif. Intell. Res. Develop., 16th Int. Conf. Catalan Assoc. Artif. Intell., vol. 256, Oct. 2013, p. 157.
- [7] L. J. S. M. Alberts. Churn prediction in the mobile telecommunications industry, chapter 2. Master’s thesis, Maastricht University, 2006.
- [8] E. Shaaban et al., "A proposed churn prediction model", *International Journal of Engineering Research and Applications*, vol. 2.4, pp. 693-697, 2012.
- [9] Sharma, A., Panigrahi, D., & Kumar, P. (2013). A neural network based approach for predicting customer churn in cellular network services. *arXiv preprint arXiv:1309.3945*.
- [10] Hung, S. Y., Yen, D. C., & Wang, H. Y. (2006). Applying data mining to telecom churn management. *Expert Systems with Applications*, 31(3), 515-524.