Telecom churn prediction using Artificial Neural Network(ANN)

Deep learning mini project

December-2021

Department of computer science engineering

Indian Institute of Information Technology(IIIT), Lucknow

Submitted by:

Rupali Patole	Aman Mehta	Shreya Goswami
MCS21005	MCS21011	MCS21023
mcs21005@iiitl.ac.in	mcs21011@iiitl.ac.in	mcs21023@iiitl.ac.in

Acknowledgement

We would like to place on record our deep sense of gratitude to DR. CHANDRANATH ADAK, Assistant Professor CSE, Indian Institute of Information Technology(IIIT), Lucknow for his generous guidance, help and useful suggestions.

He was extremely supportive and helped us in every way possible. His insightful comments and constructive suggestions helped us improve the quality of this project work.

We also wish to extend our thanks to TA's Ananya and Rohan along with developers of Stack-Overow and video tutorials available for deep learning for their immensely helpful work including open source resources, which give us inspirational video, which gave us our inspirational ideas and provide with differerent ways to overcome difficult issues.

Table of content

Introduction

The Telecom industry is one of the largest industries in every country but there is huge competition in this industry. Each company has millions of users as of now but it is predicted that the users in the telecom industry will get increase rapidly. Since the cost of getting a new customer is higher than the retention cost of an existing customer so to get better detail about customers the telecom companies need to analyses their data and need to apply some soft computing techniques to predict which customers may leave their services, this process of predicting the customers that will be leaving the firms service is known as churn prediction analysis. Churn analysis is very important as it gives more insights about the customers, the features impacting the churn rate, and using these companies can trigger different marketing strategies like a discount offer, price improvement, customer segregation, survey collection from persons having a bad experience, etc to reduce the churn rate. Churn analysis needs to be done quarterly so that company can fine-tune their offering and hence can improve quarter results by improving their customer base retention and by attracting new customers by their quarter improvisation.

Motivation and objective

Motivation**:** The Telecom industry is one of the largest industries in every country and solving any problem directly or indirectly to it will be of huge use. As the churn prediction has a great industrial use but as there is no such model that does the same with high accuracy we get a great use case to work upon and it motivates us as if we make a good model then we may get good recognition.

Objective: We are not making a general model for churn prediction for every industry instead we have targeted the telecom industry as it is one of the largest industry so we have made a model upon it and we may scale/generalize it for another use case in the near future. The objective of our model is to predict whether the customer will churn out of services or not given the customer detail; all the data that is fed into the model is from the telecom industry.

As the churn prediction is a classification problem i.e the output label is categorical so we are targeting that our model has high accuracy with good recall, precision, and F1-score.

Brief survey / Related Work

In the study by Ismail, et al on churn prediction of Malaysia's largest telecommunication provider they compare two machine learning models that are logistic regression and multi-layer perceptron. For the model evaluation, they used accuracy, sensitivity, and specificity. They trained and tested both models on different hyperparameters and found that multilayer perceptron outperforms logistic regression. While MLP gave 95.59% accuracy, 94.87% sensitivity, and 96.55% specificity, the logistic regression only gave 86.96% accuracy, 92.31% sensitivity, and 80% specificity.

In the study by Brandusoiu, et al on churn prediction analysis of prepaid telecom service providers they gave an overview of how a multi-layer perceptron is modelled for a use case. They used accuracy for model their model evaluation but they also discussed various other ways of model evaluation. Their model got 99.55% accuracy.

In the study by Adwan, et al on churn prediction analysis of telecom service providers of Jordan they modelled a multi-layer perceptron. They compared the performance of the model at different hyperparameter tuning. For evaluating the performance they used accuracy, hit rate and churn rate. They also did impact factor analysis to study the impact of each feature on the target label i.e on the churn.

Dataset details

For making a good model we need good data from some reliable resource so to make our model we used the dataset provided by the Kaggle which is the most reliable community platform which provides good quality dataset.

Dataset URL: https://www.kaggle.com/blastchar/telco-customer-churn

Dataset details:

Dataset characteristic	Multivariate, Text
Attribute characteristics	Real
Associated task	Classification
Number of samples	7043
Number of features	21
Missing values	Yes

Model Detail

For meeting our objective we made an artificial neural network(ANN) based model but as the data were having some missing values and some textual data we did feature engineering before feeding it to our model. The flow of our process/model diagram is given below.

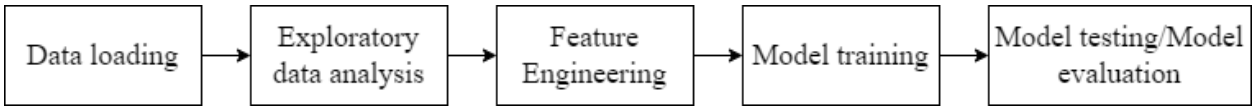

Fig: Model Diagram

The details of each step are as follows:

1. Data loading: The dataset was structured and data was present in CSV format so we used the pandas framework of python to load the data.
2. Exploratory data analysis: After loading the data we did EDA to get to know more about our data. In this, we plotted different plots with our label like histogram plot on monthly charges, sunburst plot with our label along with tech support and tenure features, etc.
3. Feature engineering: Our data was having some missing values and was having some textual data so we did feature engineering before feeding the data to our model.
   Feature TotalCharges was having null values so to deal with it we delete the samples that were having null value, we did this as there was very less samples that were having null values in TotalCharages feature.
   Few features were having categorical data but the data was textual so to replace the textual data we replace them with the category count eg Churn feature was having values Yes and No so we replace Yes with 1 and No with 0.

4. Model training: Post feature engineering we spitted our data into training and testing data. We kept 20% samples for testing and 80% for training. We made our ANN model using Keras, our model was having three layers first layer was having 64 neurons next one was having eight neurons and the last one was having one neuron. We used relu as an activation function for the first two layers and sigmoid in our output layer. We used l2-regularization along with 20% dropout at the first two layers. We used binary cross-entropy loss as the loss function and trained our model upon 150 epochs with a learning rate decay of 0.001 at every 50 epochs.
5. Model testing/model evaluation: For evaluating the performance of our model we used the Accuracy, precision, recall and F1-score as evaluation metrics.

Experiment and results

We trained our model at 150 epochs the results that our model produced is below.

Parameter	Score
Accuracy	0.80
Precision(macro average)	0.73
Precision(weighted average)	0.81
Recall(macro average)	0.76
Recall(weighted average)	0.80
F1-score(macro average)	0.74
F1-score(weighted average)	0.80

Conclusion

There were very few models available for the churn prediction. We made a simple yet effective model that has performed pretty well on model evaluation parameters. Results produced was good but in the future model can be improved by employing new feature engineering techniques, increasing the epochs, changing the hyper-parameters, etc.

References

[1] Ismail, Mohammad Ridwan, et al. "A multi-layer perceptron approach for customer churn prediction."

[2] CHURN PREDICTION IN THE TELECOMMUNICATIONS SECTOR USING NEURAL NETWORKS

[3] Predicting Customer Churn in Telecom Industry using Multilayer Preceptron Neural Networks: Modeling and Analysis

Papers:

[1] Ismail, Mohammad Ridwan, et al. "A multi-layer perceptron approach for customer churn prediction."

[2] CHURN PREDICTION IN THE TELECOMMUNICATIONS SECTOR USING NEURAL NETWORKS

[3] Predicting Customer Churn in Telecom Industry using Multilayer Preceptron Neural Networks: Modeling and Analysis

Web links:

https://www.kaggle.com/
https://stackoverflow.com/
https://www.youtube.com/
https://keras.io/
https://d2l.ai/