Customer Churn Prediction

Dr.K.Geetha
Assistant ProfessorComputerScienceandEngineering Department

SRM Institute of Science and TechnologyKattankulathur, Tamil Nadu, Indiageethak5@srmist.edu.in

Prachi Tomar
Computer Science Engineering
SRM Institute of Science and Technology
Chennai, India.
line5:pt5353@srmist.edu.in

Anisha Jain
Computer Science Engineering
SRM Institute of Science and Technology
Chennai, India.
line5:aj7552@srmkist.edu.in

***Abstract*—With the rapid advancement of digital systems and related information technologies, there is a growing tendency in the global economy to develop digital CRM systems. This research uses a real-world study to predict customer churn and recommends the use of PyCaret to improve a customer churn prediction model. Unlike most studies, this work seeks to employ PyCaret Toolkit, an open source machine learning library designed to make executing typical activities in a machine learning project simple.**

As a result, a client cluster with a higher risk of fraud has been identified.

Keywords— Churn, PyCaret, CRM.

1.INTRODUCTION

1.1 The need for Customer Churn Rate Prediction

In a retail sector business, some customers stick around while others stop shopping at a particular store after certain period of time. Detecting which customers have decided to shop elsewhere and which ones are idle at the moment, is a Herculean's task for a company. Customer churn is the tendency of customers to stop purchasing with a company over a time period. Customer churn is also called customer attrition or customer defection. Churning impedes growth. Therefore, companies should have a proper defined method to compute customer churn rate for a given time. By keeping track of churn rate, organizations can be equipped to success rate in terms of customer retention.

A. Objectives

The main objectives of this project are listed below:

- To predict churn value for all the customers of the company for a given period of time.
- To compute the overall churn rate for the given time.
- To provide a deeper insight into the sales by analyzing customers' buying pattern.
- To detect customers who are about to drop out from the business in order to take necessary steps.
- To provide clear visualizations of the churn predictions to help businesses come up with better strategies.
- To help businesses know the real value of a potential churn customer and retain him/her as a loyal customer by establishing priorities, optimizing resources, putting efficient business efforts and maximizing the value of the portfolio of the customer. • To help businesses come up with personalized customer retention plans to reduce the churn rate.

2.LITERATURESURVEY

A. *The authors of Adbelrahim et al. used tree-based algorithms to forecast customer turnover, including decision trees, random forests, GBM tree methods, and XGBoost. In a comparison, XGBoost outperformed the competition in terms of AUC accuracy. The precision of the feature selection process can, however, be increased further by employing optimization methods. Support vector machine, decision tree, naive bayes, and logistic regression were used in a comparative comparison of machine learning models for customer churn prediction by Praveen et al.*

B. *In a CRM dataset provided by American telecom companies, B.Chih-Fong Tsai proposed hybrid neural networks algorithms to forecast customer churners. They created two hybrid models for churn prediction by merging two different neural network methodologies, such as back-propagation artificial neural networks (ANN) and self-organizing maps (SOM).*

C. *Ning Lu proposed using boosting algorithms to improve a customer churn prediction model in which customers are divided into two clusters based on the boosting algorithm's weight. As a result, a potentially dangerous customer cluster has been discovered. As a basis learner, logistic regression is employed, and a churn prediction model is developed for each cluster separately. When compared to a single logistic regression model, the experimental findings indicated that the boosting approach gives a good separation of churn data.*

III PROPOSED SYSTEM

3.1 Overview:

The major goal of the customer churn forecasting model is to proactively engage with customers who are most likely to churn. Consider the following scenario: To increase their lifetime worth to the company, give them a gift voucher or any promotional pricing and lock them in for another year or two.

There are two major things to grasp in this situation:

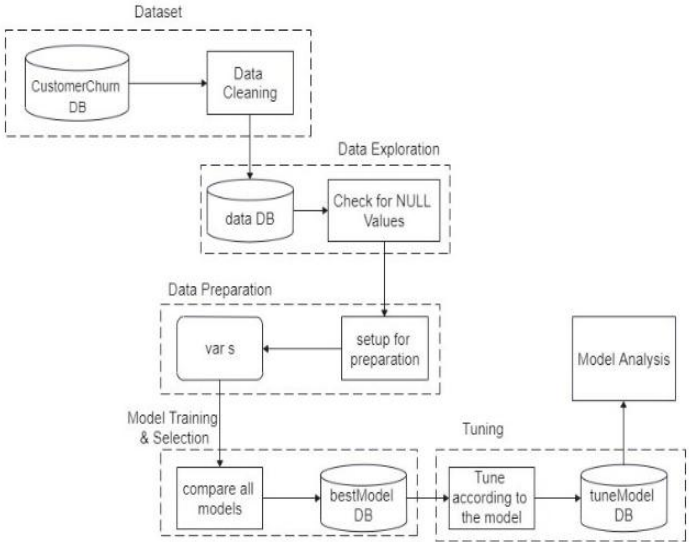
We want a customer churn predictive model that can predict attrition in advance (say, one month, three months, or even six months ahead of time – it all depends on the use-case). This implies you must be very careful about the cut-off date, and you should not use any information after the cut-off date as a feature in the machine learning model, otherwise it will be leakage. The Event is the time period preceding the cut-off date.

In most cases, you'll have to work a little bit to build a target column for customer churn prediction; it's usually not available in the form you'd like. If you want to anticipate whether a client would churn in the next quarter, for example, you'll go through all of your active customers as of your event cut-off date and see if they left the company in the next quarter or not (1 for yes, 0 for no). Performance Window is the name of the quarter in this situation.

3.PROPOSEDWORK

This machine learning model will be employed in the business knowing how the data is gathered and the churn target is produced (which is one of the more difficult components of the challenge). From left to right, as seen in the fig 3.1

- Customer churn history is fed into a model (event period for X features and performance window for target variable).
- Every month, the active customer base is fed into a Machine Learning Predictive Model to calculate the likelihood of each client churning (in business lingo, this is sometimes called a score of churn).
- The list will be sorted from highest to lowest probability value (or score), and customer retention teams will begin engaging with customers to prevent churn, usually by offering a promotion or gift card to lock in a few more years.
- Clients with a low likelihood of turnover (or, in other words, customers for whom the model forecasts no churn) are satisfied customers. They are not dealt with in ny way.



3.1 ArchitectureDiagram

4. IMPLEMENTATION

Dataset

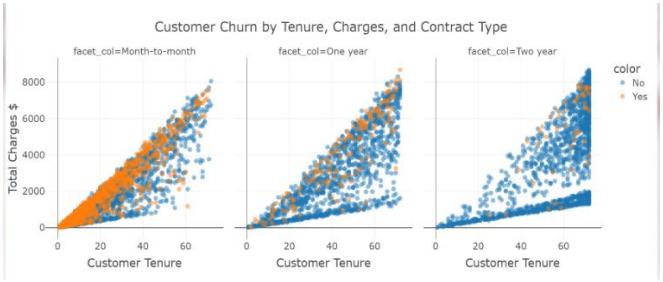
Kaggle dataset called Telecom Customer Churn is used for this experiment. The target column is already there in the dataset, so we may utilise it as is.

Exploratory Data Analysis

It's worth noting that TotalCharges is an object rather than a float64. After further study, discovered that this column contains some blank spaces, causing Python to force the data type to object. We'll have to trim blank spaces before altering the data type to fix this.

When it comes to customer churn or retention, contract type, tenure (duration of customer stay), and price strategies are all crucial pieces of information.

The contracts marked "Month-to-Month" have the highest rate of churn. Of course, that's logical. Also, when tenure and total charges increase, the likelihood of consumers with high tenure and low charges is lower than customers with high tenure and high costs.



4.1 Scatter Plot

Missing Values

Because we replaced blank values with np.nan, Total_Charges now has 11 rows with missing data.leave it to PyCaret to automatically infer it.

Data Preparation

The setup is the first and only obligatory step in any machine learning experiment performed in PyCaret, and it is common to all modules in PyCaret. This function handles all of the data preparation that must be done before training models. PyCaret supports a wide range of pre-processing tools in addition to completing some simple default processing chores.

To learn more about PyCaret's preprocessing features.
As a result, it will generate forecasts there is no need to join IDs manually.

Model Training & Selection

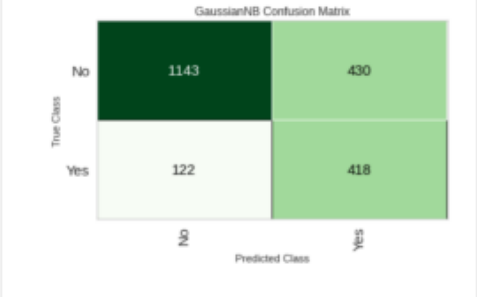
The training process starts by using the compare models functionality now that the data has been prepared. This function uses cross-validation to train all of the algorithms in the model library and to evaluate numerous performance measures. PyCaret's tune model method automatically tunes the model's hyperparameters.

```
# compare all models
best_model = compare_models(sort='AUC')
```

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	Profit	TT (Sec)
gbc	Gradient Boosting Classifier	0.8006	0.8451	0.5440	0.6800	0.6955	0.4651	0.4695	251600.0	0.692
lr	Logistic Regression	0.8014	0.8436	0.5506	0.6802	0.6997	0.4691	0.4731	254600.0	0.441
ada	Ada Boost Classifier	0.8002	0.8431	0.5365	0.6812	0.6911	0.4610	0.4662	248300.0	0.271
lda	Linear Discriminant Analysis	0.7978	0.8368	0.5613	0.6468	0.6000	0.4657	0.4684	257000.0	0.042
lightgbm	Light Gradient Boosting Machine	0.7901	0.8333	0.5236	0.6340	0.5728	0.4355	0.4395	238200.0	0.161
nb	Naive Bayes	0.7469	0.8297	0.7720	0.5216	0.6220	0.4425	0.4623	315900.0	0.026
rf	Random Forest Classifier	0.7868	0.8185	0.4861	0.6345	0.5562	0.4190	0.4264	226400.0	0.813
et	Extra Trees Classifier	0.7669	0.7896	0.4921	0.5807	0.5317	0.3782	0.3811	214200.0	0.765
knn	K Neighbors Classifier	0.7582	0.7500	0.4319	0.5686	0.4895	0.3354	0.3415	185900.0	0.139
dt	Decision Tree Classifier	0.7294	0.6616	0.5124	0.4986	0.5044	0.3187	0.3193	203800.0	0.040
qda	Quadratic Discriminant Analysis	0.5424	0.5765	0.6507	0.3274	0.4279	0.1155	0.1415	166900.0	0.027
dummy	Dummy Classifier	0.7304	0.5000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0	0.021
svm	SVM - Linear Kernel	0.7024	0.0000	0.5198	0.5578	0.4751	0.2951	0.3055	193500.0	0.059
ridge	Ridge Classifier	0.7996	0.0000	0.5244	0.6656	0.5857	0.4660	0.4623	243200.0	0.021

Model Analysis

```
[22] # confusion matrix
     plot_model(best_model, plot = 'confusion_matrix')
```



4.2 Confusion matrix

Training of numerous models to identify the best model with the highest AUC, then tweaked the best model's hyperparameters to squeeze a little more AUC out of it. The best AUC, on the other hand, does not always imply the best business strategy.
In a churn model, the benefit of true positives is frequently greater than the cost of false positives.

5. RESULTS

We were able to exclude non-essential data and forecast the model that will be used to train the data ahead of time, resulting in optimal results, i.e. lower customer turnover and more earnings, by following the processes outlined above.

6. CONCLUSION

We were able to train numerous models and select the one that matters to the business with only a few lines

of code. I'm a regular blogger who primarily covers PyCaret and its real-world applications.

7. REFERENCES

[1] Customer attrition. Retrieved from http://en.wikipedia.org/wiki/Customer_attrition, on, February 25, 2011.
[2] Predictive analytics, Retrieved from http://en.wikipedia.org/wiki/Predictive_analytics on April 14, 2011.
[3] Predictive modeling, Retrieved from http://en.wikipedia.org/wiki/Predictive_modelling, on March 17, 2011.
[4] R. Mattison, The Telco Churn Management Handbook, 2001.
[5] Churn Analysis. (n.d.). Retrieved fromhttp://www.ambarasoft.com/researchservices/churnanalysis.html
[6] I. H. Witten, E. Frank. Data Mining Practical Learning Tools and Techniques, Morgan Kaufmann Publishers, 2005.
[7] Supervised learning Retrieved from http://en.wikipedia.org/wiki/Supervised_learning, retrieved on April 16, 2011.

[8] Lemmens, A., & Croux, C., "Bagging and boosting classification trees to predict churn," DTEW Research Report, (2003). 0361.

[9] Liaw, A., & Wiener, M., "Classification and regression by random forest," The Newsletter of the R. Project, (2002). 2(3), 18–22.

[10] "Introduction to Support Vector Machines ----- Opencv 2.4.13.2 Documentation". *Docs.opencv.org*. [Online] Available: http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html

[11] "Learn Gradient Boosting Algorithm for Better Predictions (With Codes In R)". *Analytics Vidhya*. [Online] Available: https://www.analyticsvidhya.com/blog/2015/09/completeguide boosting-methods/

[12] "Gradient Boosting"*En.wikipedia.org*. [Online] Available: https://en.wikipedia.org/wiki/Gradient_boosting.