Churn Prediction in Telecom Industries using Machine Learning Techniques

Mandar Vaidya
X17153409
MSc. in Data Analytics

Background and Introduction

- Customer churn is extensively famailiar concept in todays global world.
- Churn is defined as a customer leaving a service.
- Almost 30-40% of churn rate, telecomunication industries takes the first place on the list of churners.
- Due to fierce competition in telecom industries, companies have shifted their focus on customer retention than customer acquisition.
- That's why telecom companies are focusing more on defensive marketing startegies.
- To analysed about churners and their behavior various data mining techniques are implemeted for the prediction.
- There are various reasons for churn, to study this reasons and patterns, machine learning algorithms are implemented.

Research Question:

- How will organizations benefit to cut down churners and churn rate from predictive model with the help of Machine Learning algorithm

Methodology

- Since, sequential algorithms do not store and process all the previous data, it is more efficient for training and testing data, although real-life scenario might be different
- SVM has a better accuracy but it can't run on real-time as well as big data.
- Some KPIs have unique behavioral patterns with some unique features.
- So, while Machine Learning algorithms it affects dimensionality reduction thus by causing problems in Incremental learning
- SVM kernels like sigmoid, polynomial, linear with the help of RBF can perform better but accuracy is low thus different approaches with different predictors can be used.
- Churn column is the key column as it shows which customers have left the company.
- Churn is a key attribute. The customers in the dataset are classified as dichotomous variable called Churn (either yes/no).
- A customer will be categorized as Active (Yes) if he/she continues to use the same network (non-churner). Whereas, a customer will be categorized as Non-Active (No) of if he/she discontinues the network (churner).

Methodology

- Determining the Important Variables
- Determining the Correlation between the Variables
- Checking about Class Imbalance
- Logistic Regression
- K-means Clustering
- Random Forest
- Singular Value Decomposition

Conclusion

- Thus, by using all the above alogrithim, churn rate and percentage of churn can be calcualted for better performance and thus by helping organisation in customer retention than acquisition.

Video Link:

https://youtu.be/ENuIrIncxXM

Thank You ☺