

Telecom-Churn-Case-Study

By: Ayushi and Akash

Random Forest, XgBoost, Logistic Regression

In the telecom industry, customers are able to choose from multiple service providers and actively switch from one operator to another. In this highly competitive market, the telecommunications industry experiences an average of 15-25% annual churn rate. Given the fact that it costs 5-10 times more to acquire a new customer than to retain an existing one, customer retention has now become even more important than customer acquisition.

To reduce customer churn, telecom companies need to predict which customers are at high risk of churn.

In this project, customer-level data of a leading telecom firm is analysed, predictive models are built to identify customers at high risk of churn and the main indicators of churn are identified.

The dataset contains customer-level information for a span of four consecutive months - June, July, August and September. The months are encoded as 6, 7, 8 and 9, respectively.

The business objective is to predict the churn in the last (i.e. the ninth) month using the data (features) from the first three months.

Data Analysis Understanding the typical customer behaviour during churn will be helpful as customers usually do not decide to switch to another competitor instantly, but rather over a period of time (this is especially applicable to high-value customers). In churn prediction, we assume that there are three phases of customer lifecycle :

The 'good' phase: In this phase, the customer is happy with the service and behaves as usual.

The 'action' phase: The customer experience starts to sore in this phase, for e.g. he/she gets a compelling offer from a competitor, faces unjust charges, becomes unhappy with service quality etc. In this phase, the customer usually shows different behaviour than the 'good' months. Also, it is crucial to identify high-churn-risk customers in this phase,

since some corrective actions can be taken at this point (such as matching the competitor's offer/improving the service quality etc.)

The 'churn' phase: In this phase, the customer is said to have churned. Churn is defined based on this phase. Also, it is important to note that at the time of prediction (i.e. the action months), this data is not available for prediction. Thus, after tagging churn as 1/0 based on this phase, all data corresponding to this phase are discarded.

In this case, since data over a four-month window is analysed, the first two months are the 'good' phase, the third month is the 'action' phase, while the fourth month is the 'churn' phase.