

Conventional statistical methods (e.g. logistics regression, decision tree, and etc.) are very successful in predicting customer churn. However, these methods could hardly predict when customers will churn, or how long the customers will stay with. The goal of this study is to apply survival analysis techniques to predict customer churn by using data from a telecommunications company.

This study will help telecommunications companies understand customer churn risk and customer churn hazard in a timing manner by predicting which customer will churn and when they will churn. The findings from this study are helpful for telecommunications companies to optimize their customer retention and/or treatment resources in their churn reduction efforts.

0.1 Churn

In the telecommunications industry, the broad definition of churn is the action that a customer's telecommunications service is canceled. This includes both service-provider initiated churn and customer initiated churn. An example of service-provider initiated churn is a customer's account being closed because of payment default. Customer initiated churn is more complicated and the reasons behind vary. In this study, only customer initiated churn is considered and it is defined by a series of cancel reason codes. Examples of reason codes are: unacceptable call quality, more favorable competitor's pricing plan, misinformation given by sales, customer expectation not met, billing problem, moving, change in business, and so on.