

Predicting Customer Churn – Telecoms

Nihal, James, Sahas, Laine, Tyler

Today, the telecom industry has a market cap of $2.6 trillion globally. This is a massive industry in a highly competitive market where the customer base is very aware of the alternatives available to them. Making customers churn all but a certain – in fact, 14% of US respondents answered on a survey on “Willingness to change mobile carrier” with “very likely”.

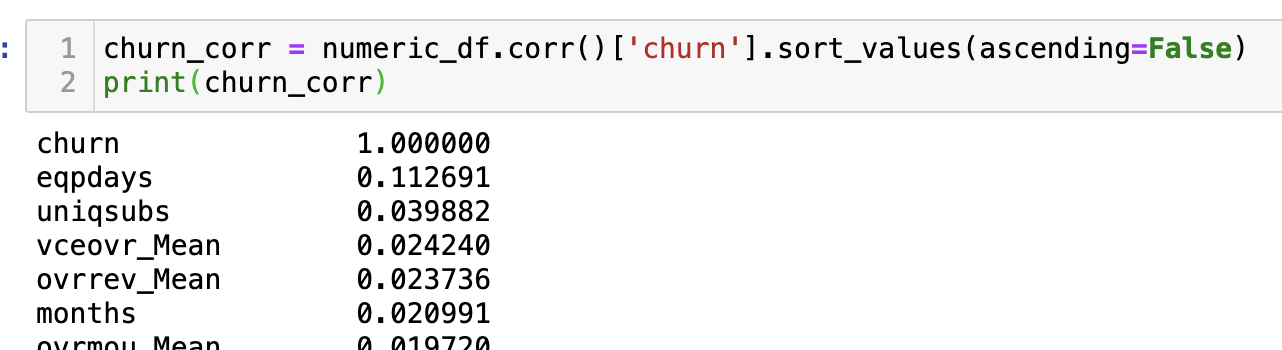
So, we know there will be some churn. The question is - Can we predict which of these customers is likely to churn based on data, giving us a chance to reduce this risk?

Our [data set](https://www.kaggle.com/datasets/abhinav89/telecom-customer/data) to help us answer this consists of 100 thousand records (each indicative of a single telecom customer) and 100 attributes related to each of those customer’s. These attributes are things like usage, revenue, call behavior, demographics, and equipment (phone) details.

Our primary target variable is **churn**, which indicates whether a customer left the service within 31-60 days after the observation date. Here is a high-level summary of the data we’re working with:

* **Usage Metrics** (e.g., mou\_Mean, totmou, attempt\_Mean, complete\_Mean)
* **Revenue Metrics** (e.g., rev\_Mean, totrev, avgrev, adjrev)
* **Call Performance & Quality** (e.g., drop\_vce\_Mean, drop\_blk\_Mean, unan\_vce\_Mean, blck\_vce\_Mean)
* **Customer Service Interaction** (e.g., custcare\_Mean, cc\_mou\_Mean)
* **Demographics** (e.g., income, ethnic, marital, numbcars, dwllsize)
* **Handset & Account Information** (e.g., hnd\_price, eqpdays, asl\_flag, phones)
* **Recent Behavioral Trends** (e.g., change\_mou, change\_rev, avg3mou, avg6rev)

In our early exploratory data analysis we found that:



A black and white text

AI-generated content may be incorrect.

✅ Most Positively Correlated Feature eqpdays (r = +0.11) → Slightly higher churn as equipment days increase (older phones are more likely to churn)

❌ Most Negatively Correlated Features hnd\_price (r = –0.10) → Lower churn among customers with more expensive handsets

From our base model we found  
A screenshot of a phone

AI-generated content may be incorrect.

|  |  |
| --- | --- |
| change\_mou | Percentage change in monthly minutes of use vs previous three month average |
| mou\_Mean | Mean number of monthly minutes of use |
| totmrc\_Mean | Mean total monthly recurring charge |
| months | Total number of months in service |
| change\_rev | Percentage change in monthly revenue vs previous three month average |
| eqpdays | Number of days (age) of current equipment |