Predicting customer Churn and Analysis

What is churn?

Customer Churn:-

It is when an existing customer, user, subscriber, or any kind of return client stops doing business or ends the relationship with company.

Analysis of churn:-

Churn is a good indicator of growth potential. Churn rates track lost customers, and growth rates track new customers—comparing and analyzing both of these metrics tells you exactly how much your business is growing over time. If growth is higher than churn, you can say your business is growing. If churn is higher than growth, your business is getting smaller.

What is churn rate?

the churn rate (also called attrition rate) measures the number of individuals or items moving out of a collective group over a specific period

Churn rate mathematical formula:-

$$\text{Churn rate} = \frac{\text{Number of customers lost during a time frame}}{\text{Number of customers at the beginning of the time frame}}$$

.

Steps to analysis of customer churn prediction:

- Identifying at-risk customers
- Identifying customer pain points,
- Identifying strategy/methods to lower churn and increase customer retention

I found some of problems to churn model as follows:

- Inaccurate or messy customer data,
- Weak attrition exploratory analysis,
- Lack of information and domain knowledge,
- Lack of a coherent selection of a suitable churn modeling approach,
- Choice of metrics to validate churn model performance,
- Line of business (LoB) of services or products,
- Churn event censorship,
- Concept drift based on changes in customers behaviour patterns driving churn,
- Imbalance data (class imbalance issue).

Churn prediction use cases on real world application:

- Telecommunication (cable or wireless network segment),
- Software as a service provider (SaaS),
- Retail market,
- Subscription-based businesses (media, music and video streaming services, etc.),
- Financial institutions (banking, insurance companies, Mortgage Companies, etc.),
- Marketing,
- Human Resource Management (Employee turnover).

I predict Telecom churn prediction system use case and analyze it:

According to define churn rate in the above it helps very important in the telecommunications industry (wireless and cable service providers, satellite television providers, internet providers, etc). The churn rate in this use case provides clarity on the quality of the business, shows customer satisfaction with the product or service, and allows for comparison with competitors to gauge an acceptable level of churn.

About the dataset

The sample data tracks a fictional telecommunications company, Telco. It's customer churn data sourced by the IBM Developer Platform, and it's

available here. It includes a target label indicating whether or not the customer left within the last month, and other dependent features that cover demographics, services that each customer has signed up for, and customer account information. It has data for 7043 clients, with 20 features.

The dataset has 7043 rows and 21 columns.

There are 17 categorical features:

- **CustomerID:** Customer ID unique for each customer
- **gender:** Whether the customer is a male or a female
- **SeniorCitizen:** Whether the customer is a senior citizen or not (1, 0)
- **Partner:** Whether the customer has a partner or not (Yes, No)
- **Dependent:** Whether the customer has dependents or not (Yes, No)
- **PhoneService:** Whether the customer has a phone service or not (Yes, No)
- **MultipeLines:** Whether the customer has multiple lines or not (Yes, No, No phone service)
- **InternetService:** Customer's internet service provider (DSL, Fiber optic, No)
- **OnlineSecurity:** Whether the customer has online security or not (Yes, No, No internet service)

- **OnlineBackup:** Whether the customer has an online backup or not (Yes, No, No internet service)
- **DeviceProtection:** Whether the customer has device protection or not (Yes, No, No internet service)
- **TechSupport:** Whether the customer has tech support or not (Yes, No, No internet service)
- **StreamingTV:** Whether the customer has streaming TV or not (Yes, No, No internet service)
- **StreamingMovies:** Whether the customer has streaming movies or not (Yes, No, No internet service)
- **Contract:** The contract term of the customer (Month-to-month, One year, Two years)
- **PaperlessBilling:** The contract term of the customer (Month-to-month, One year, Two years)
- **PaymentMethod:** The customer's payment method (Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic))

Next, there are 3 numerical features:

- **Tenure:** Number of months the customer has stayed with the company
- **MonthlyCharges:** The amount charged to the customer monthly

- **TotalCharges:** The total amount charged to the customer

Finally, there's a prediction feature:

- **Churn:** Whether the customer churned or not (Yes or No)

These features can also be subdivided into:

- Demographic customer information:
- gender , SeniorCitizen , Partner , Dependents
- Services that each customer has signed up for:
- PhoneService , MultipleLines , InternetService , OnlineSecurity , OnlineBackup , DeviceProtection , TechSupport , StreamingTV , StreamingMovies,
- **Customer account information:**
- tenure , Contract , PaperlessBilling , PaymentMethod , MonthlyCharges , TotalCharges

We can train the data and anlyze the prediction on churn:

1. By using matplotlib library we can calculate churn percentage of assumed dataset.according to our choose dataset contains

- Churn: No – 73.5%

- Churn: Yes – 26.5%

Without get in percentage we want to get numbers wise of churn we use value_counts method.

2. we churn the date according to monthly charge and phone service by using boxplot.

3. Then train the data of telecommunication industry and it gives a detail information about howmany churn in their company or not and their expenses after it could be increase their income is stable or not.

S180187

S180390
