

Telecom Customer Churn Prediction

Abstract

The aim of this project is to develop an application for a telecommunication company's marketing, product development, and sales teams. The application utilizes customer churn data to provide valuable insights and visualizations that can guide product developments, marketing strategies, and sales approaches. By analysing customer churn, the project aims to identify patterns and trends that can help the company retain customers, enhance customer satisfaction, and reduce customer attrition.

Customer churn refers to the rate at which customers decide to discontinue their services with a company. In the telecommunications industry, customer churn represents the number of subscribers who switch to a competitor's services. It is a critical metric for businesses, as acquiring new customers is more costly than retaining existing ones. By analyzing and continuously monitoring customer churn, companies can identify areas of improvement in their customer attrition strategies. Furthermore, understanding the factors that contribute to churn can facilitate the development of customized service packages and solutions.

Approach and Motivation

The motivation behind this project stems from the significance of customer churn analysis in the telecommunications industry. Customer churn has a substantial impact on a company's revenue and profitability, making it more cost-effective to retain existing customers than acquire new ones. By understanding the reasons behind churn and proactively addressing customer concerns, companies can reduce churn rates, foster customer loyalty, and drive business growth.

The developed application will enable data-driven decision-making and strategy formulation. It will empower the marketing team to identify high-risk customer segments and design personalized retention campaigns to maximize effectiveness and improve customer retention rates. The application will also provide insights to the product development team, helping them identify areas of improvement and prioritize enhancements that align with customer expectations, leading to increased satisfaction and loyalty.

Moreover, the application will benefit the sales team by providing valuable information for customer interactions. Understanding the factors influencing customer satisfaction and churn will enable the sales team to recommend suitable products, address concerns promptly, and nurture long-term relationships. This customer-centric sales approach will enhance retention and drive revenue growth through upselling and cross-selling opportunities.

In summary, this project's approach of utilizing customer churn data and developing an application aligns with the company's goals of improving customer retention, enhancing satisfaction, and achieving business success. By harnessing the power of data analytics, the telecommunication company can gain a competitive advantage, optimize operations, and cultivate enduring customer relationships.

Dataset

For this project, a dataset containing 7043 observations (rows) and 22 variables (columns), having information about customer demographics (gender, senior citizenship status, children, and marital status), services they signed up for (phone line, multiple lines, online security, online backup, device protection, tech support, streaming TV and Movies), account information (type of contract, payment type, paperless billing, monthly and total charges), and churn (which customers left within the past month when the data was collected). The dataset has been pre-processed leading to the creation of a new dataset `preprocessed_data.csv`.

Pre-Processing

The app uses Pandas and NumPy library for pre-processing data, plotly, seaborn and matplotlib libraries for data visualization. Not much pre-processing was required for this dataset. It required changing the data type of “TotalCharges” column from object to float which led to the finding of 11 missing values. After dropping those 11 values, 7032 observations (rows) and 22 variables (columns) were left. Also, the labels for “SeniorCitizen” were changed to Yes and No from 0 and 1 to avoid any error or warning during the app deployment.

After all the above processes were done, a new csv file was created namely `preprocessed_data.csv`. This is the final document that was used for analysing the data and building the application.

Data Analysis and Visualization

3.1 Raw Data

The Streamlit app provides an option to view the raw data. Users can toggle the "Show raw data" checkbox to see the entire dataset. Also, an input box has been provided to input the Customer Id, based on which all the data related to that particular ID will be generated in the main box.

3.2 Demographics Analysis

The app allows users to analyse customer churn based on demographics such as gender, senior citizen status, partner, dependents, and churn status. It provides interactive visualizations in the form of histograms and pie charts to show the count of customers based on each demographic feature. The choice for visualisation can be selected in the sidebar. Two different counts were created – one without churn data and one involving churn data.

Number of Customers by gender

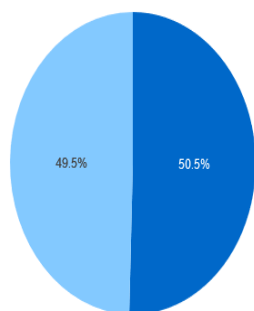


Fig. Pie Chart for gender count

Number of Customers by SeniorCitizen

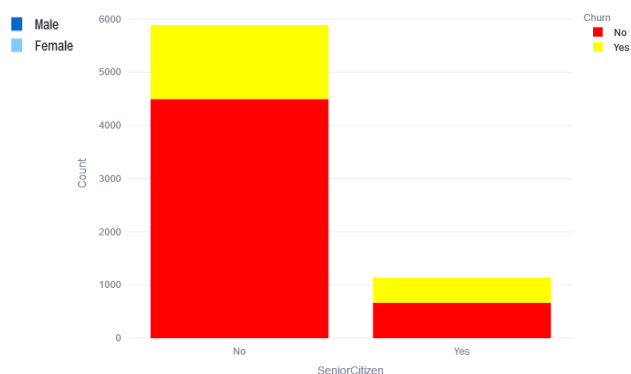


Fig. Count of Senior Citizens based on Churn.

Also, a table has been created where the total count of people and the percentage of people associated with that demographic is calculated and displayed. The following points have been analysed using the visualisations-

- About half of the customers in our data set are male while the other half are female
- There are only 16% of the customers who are senior citizens. Thus, most of our customers in the data are younger people.

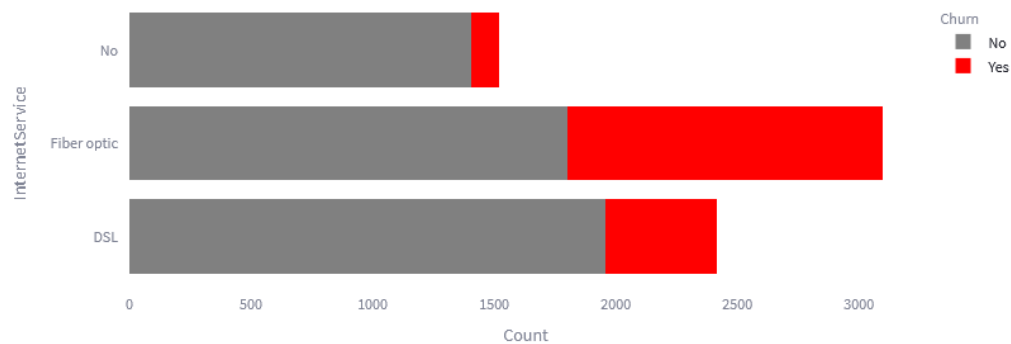
- About 50% of the customers have a partner, while only 30% of the total customers have dependents.

3.3 Service Usage Analysis

The app enables users to explore customer churn based on the services used, including phone service, multiple lines, internet service, online security, device protection, tech support, streaming TV, and streaming movies. It presents grouped bar charts to visualize the number of customers using each service, segmented by churn status.

Number of Customers by InternetService

Number of Customers by InternetService



Count of Customers by InternetService

	InternetService	Total Count	Percentage
0	DSL	2416	34.3572
1	Fiber optic	3096	44.0273
2	No	1520	21.6155

As with the demographic analysis, count table is created to display the count and percentage of the Service availed by the customer. The graph can be hidden when multiple plots are displaying on the main page, using a checkbox in the sidebar. The following points have been analysed using the visualisations-

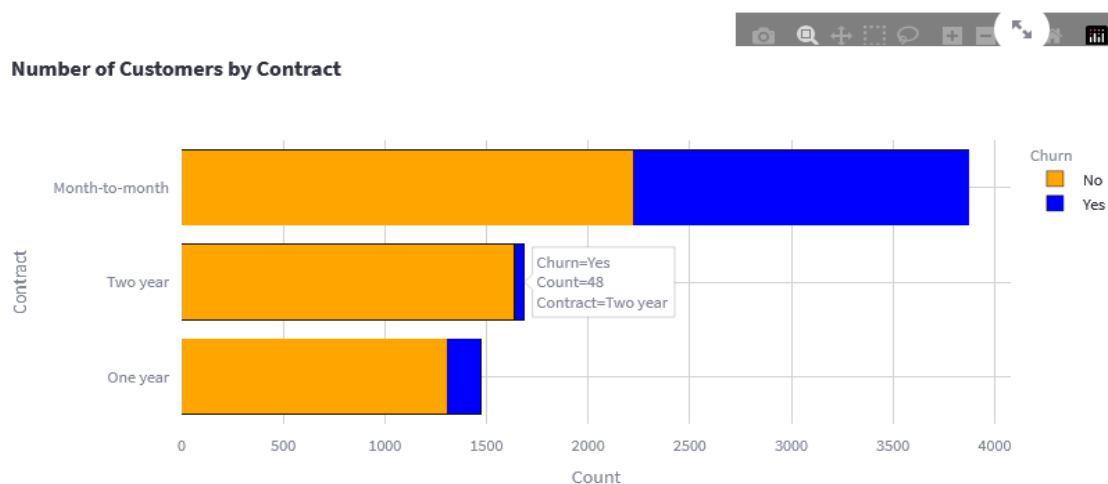
- Customers with Fiber Optic as Internet Service have left the Telecom Company in greater numbers as compared to other Internet Services.
- The same trend has been noticed with customers who have not opted for online security, tech support and device protection.

- The same trend has been observed with customers who have not opted for Streaming services for Movies and TV.

3.4 Account Information Analysis

Users can analyse customer churn based on account information, including contract type, paperless billing, and payment method. The app displays stacked bar charts to compare the number of customers in each category and their churn status.

Number of Customers by Contract



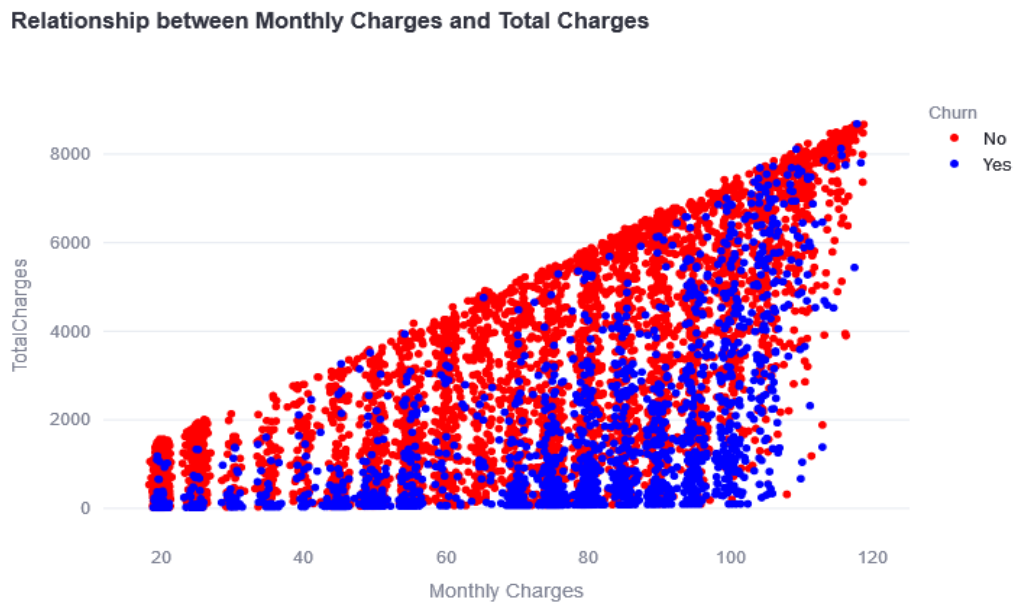
Count of Customers by Contract

	Contract	Total Count	Percentage
0	Month-to-month	3875	55.1052
1	One year	1472	20.9329
2	Two year	1685	23.9619

- As we can see from this graph most of the customers are in the month-to-month contract. While there are equal number of customers in the 1 year and 2-year contracts.
- It has been observed that the customers taking a longer contract are more loyal to the company and tend to stay with it for a longer period of time. So, the customers who have a month-to-month contract have a very high churn rate.
- Also, people paying through electronic cheques tend to leave the telecom company in greater numbers.

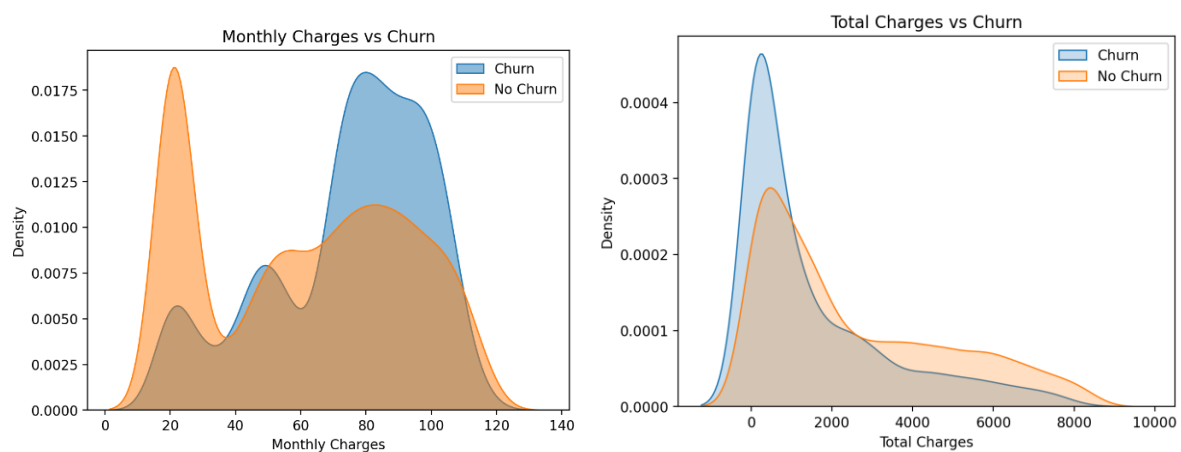
3.5 Monthly vs. Total Charges Analysis

The app provides a scatter plot to visualize the relationship between monthly charges and total charges, coloured by churn status. It helps identify any patterns or trends between charges and churn. From the Scatter plot, it is clear that the total charges increase as the monthly bill for a customer increases.



3.6 Monthly and Total Charges Analysis with Churn

The app presents KDE plots to compare the density of monthly charges and total charges for churned and non-churned customers separately using Seaborn library. This analysis helps in understanding the distribution of charges in each group.



From the graph, it is clear that-

- there is higher churn when the total charges are lower and monthly charges are high.
- there is lower churn when the monthly charges are low.

Conclusion

- The analysis has shown that the senior citizens tend to leave the telecom company in greater numbers.
- As the monthly charges increase, the churn rate also increases.
- The Streamlit app for telecom customer churn analysis provides valuable insights into customer behavior and churn patterns. By leveraging interactive visualizations, users can easily explore the data and identify factors that contribute to customer churn. This analysis can assist businesses in making informed decisions to improve customer retention and satisfaction.
- Also this application helps bridge the gap between programmers and the ease of understanding the final findings by common folks and clients.
- The application can be accessed through <https://telecom-churn.streamlit.app/>

References

1. A. Sharma, S. Mathur, and S. Goyal. (2019). Predicting Customer Churn in Telecom Industry using Machine Learning Techniques. International Journal of Computer Applications, 182(45), 8-14.
2. M. Burez and D. Van den Poel. (2009). Handling Churn in Customer Relationship Management through a Customer Segmentation Based Approach. Expert Systems with Applications, 36(3), 6127-6134.
3. V. Venkatesan, R. Kumar, and M. Ravishanker. (2007). Multiscale Modeling of Consumer Response to Direct-Mail Promotions. Journal of Marketing Research, 44(4), 640-654.
4. T. H. Davenport. (2010). Analytics at Work: Smarter Decisions, Better Results. Harvard Business Press.
5. J. Verbeke, L. Martens, and S. Baesens. (2012). Building Comprehensive Customer Profiles from Transactional Data Using Data Mining Techniques. Expert Systems with Applications, 39(8), 7068-7076.