Asmerom Fessehaye Habteab
14 January, 2021

Marketing Customer Churn Prediction Proposal

Business Question

What drives higher churn rate of customer and how can Telecoms tackle this problem?

Goal

Identify customers likely to churn and develop focused intervention plans to retain as many customers as possible.

Domain Background

Telecommunication businesses have seen the adverse effects of customer churn on company revenue. With the challenging dynamics in the telecommunications industry, reducing churn in the customer base should be among any telecommunication company's top priorities.

Companies should not only spend their valuable time and resources probing for new customers, they must keep their existing customers happy as it impacts their lifetime value and safeguards their revenue.

According to various studies, acquiring a new customer is anywhere from 5 to 25 times more expensive than retaining an existing one. Today's telecommunication companies are doing everything they can to retain their existing customers, but customer churn is still high.

Customers want great service, but surprisingly, companies aren't delivering service that exceeds or even meets their expectations which is leading to churn.

In such scenarios, what companies require is an advanced way to predict churn risk as early as possible in the customer lifecycle stage. This is possible, especially with the latest advancement and adoption in AI technologies, and such solutions can help you understand your customers' needs precisely.

Key business challenges for retaining customers

- Identifying root causes of customer dissatisfaction.
- Difficulty in identifying high-risk customers before it's too late.
- Creating personalized offers to at-risk customers.

Dataset used

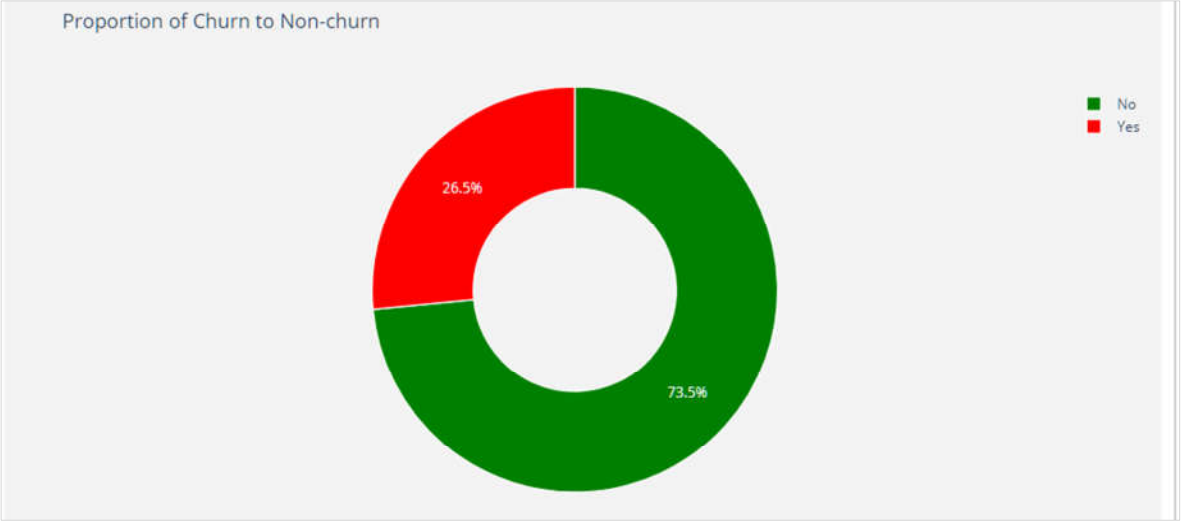
The data set includes information about:

- Customers who left within the last month – the column is called Churn
- Services that each customer has signed up for – phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies
- Customer account information – how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges
- Demographic info about customers – gender, age range, and if they have partners and dependents
- As well it consists of 7044 * 21 Cells of Info Ranging From Customer ID's to their Churn State
- Source: Telco Customer Churn programs from Kaggle

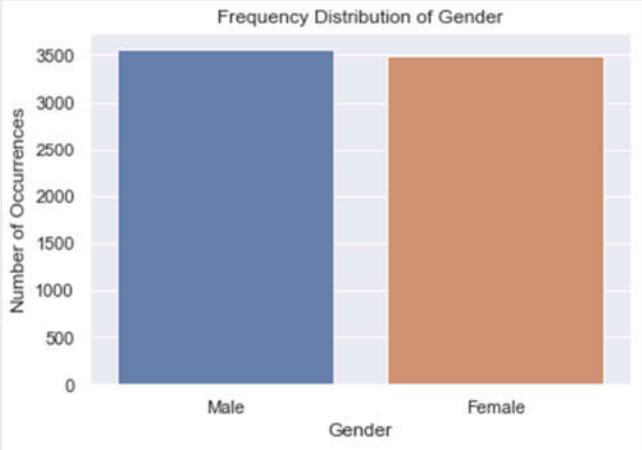
Data over View: Here are some of the reports of the data over view.

- Proportion of Churn to Non-churn
- Frequency distribution of Gender with respect to churn rate:
- Relative frequency distribution of customer as per they are senior citizen or not
- Effect of tenure period on churn rate
- Relation of Monthly Salary and Churn rate
- Relation of Customer contract type and churn rate
- Effect of Payment Method on Churn rate
- Study of effect of services on churn rate
- Relation of Fiber optics service with churn rate
- Relation of internet service subscription with online security
- Relation of online backup service with churn rate
- Relation of device protection service with churn rate
- Relation of tech support service with churn rate

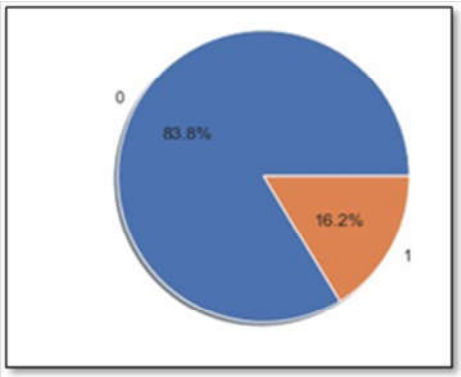
Proportion of Churn to Non-churn



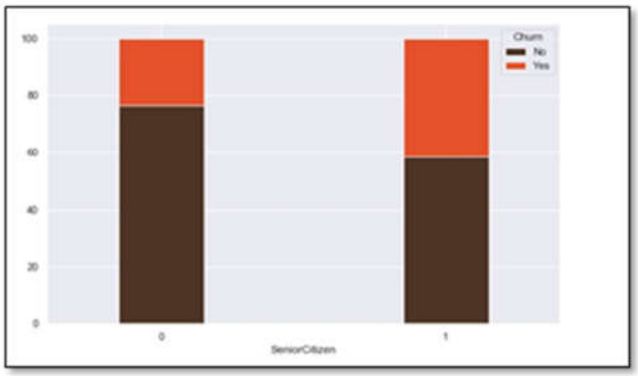
Frequency distribution of Gender with respect to churn rate:



Relative frequency distribution of customer as per they are senior citizen or not



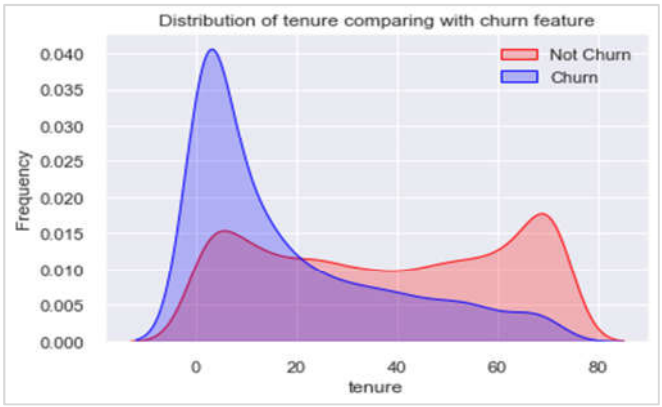
Senior citizen pie chart

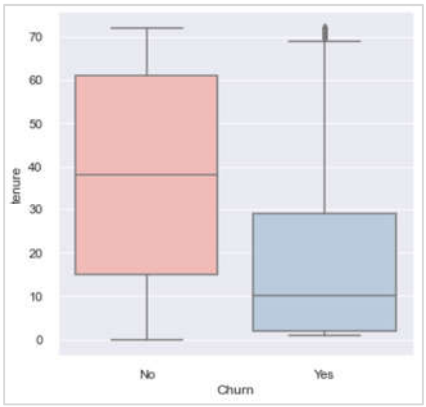


Young Citizen
Senior Citizen

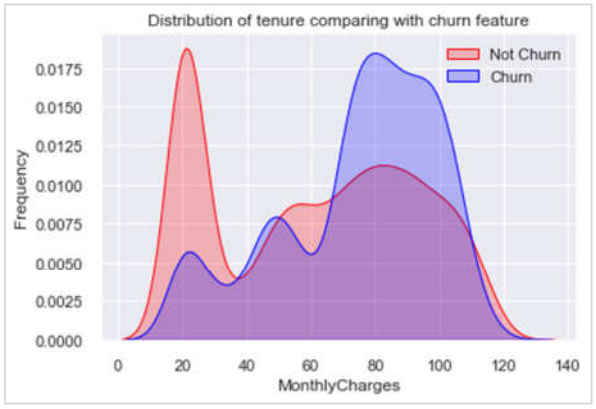
Bar chart of churn rate in

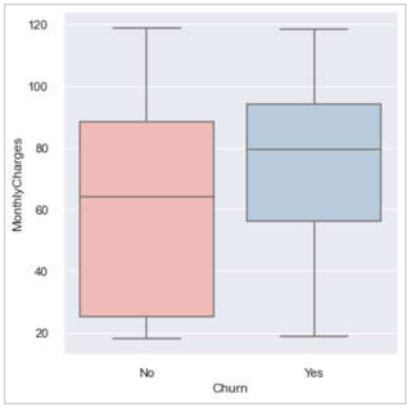
Effect of tenure period on churn rate



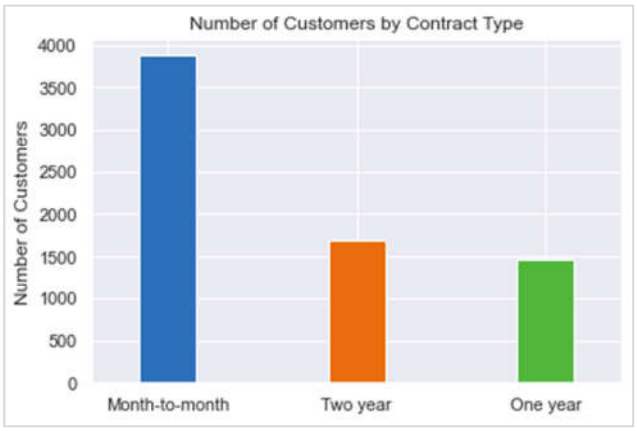


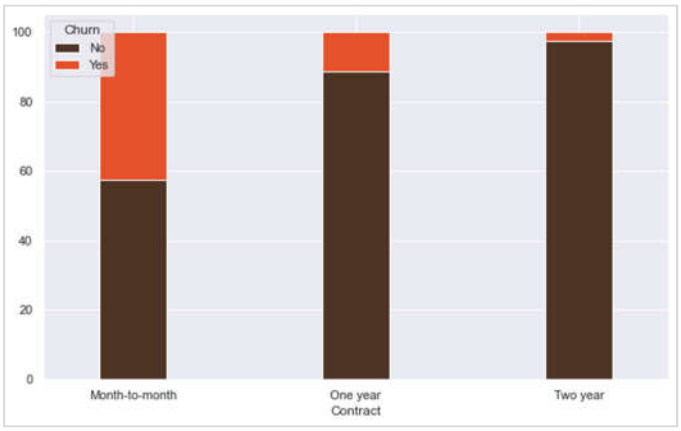
Relation of Monthly Salary and Churn rate



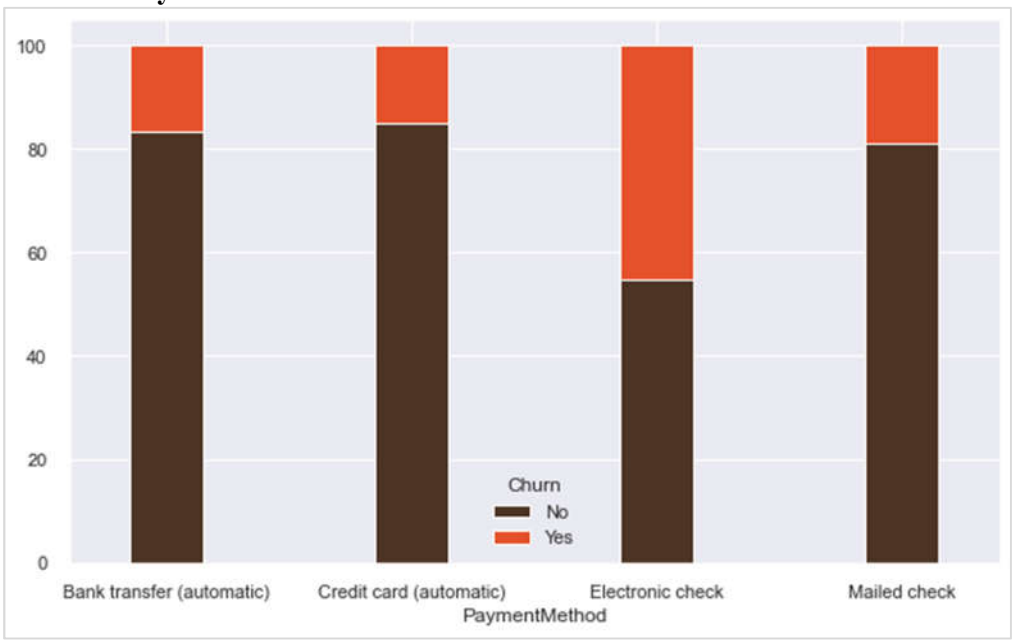


Relation of Customer contract type and churn rate

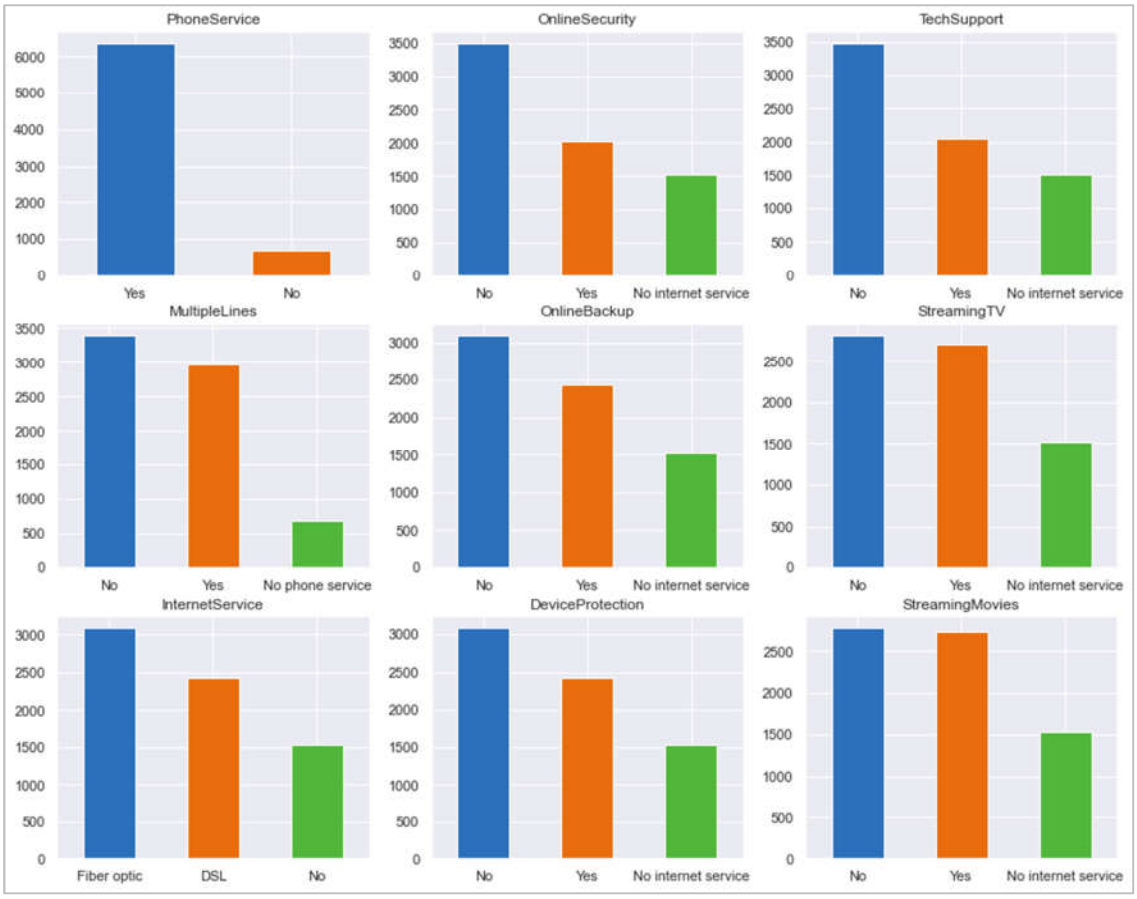




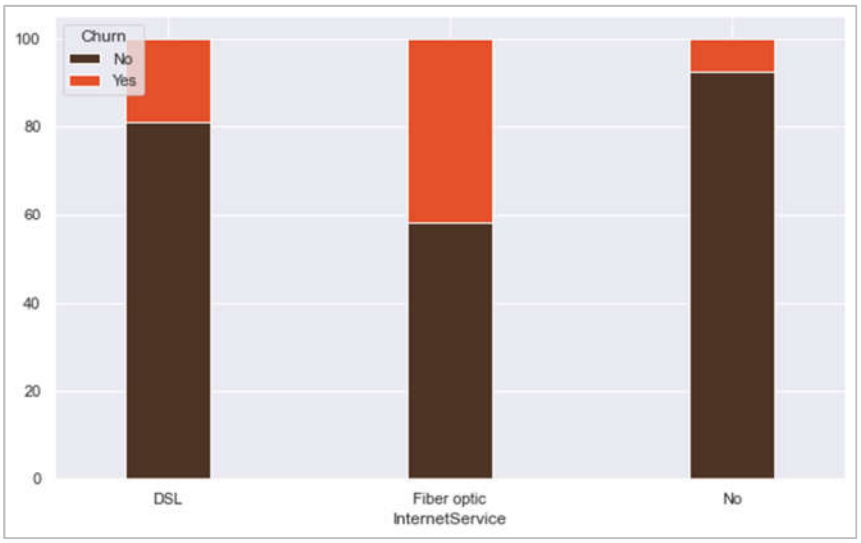
Effect of Payment Method on Churn rate



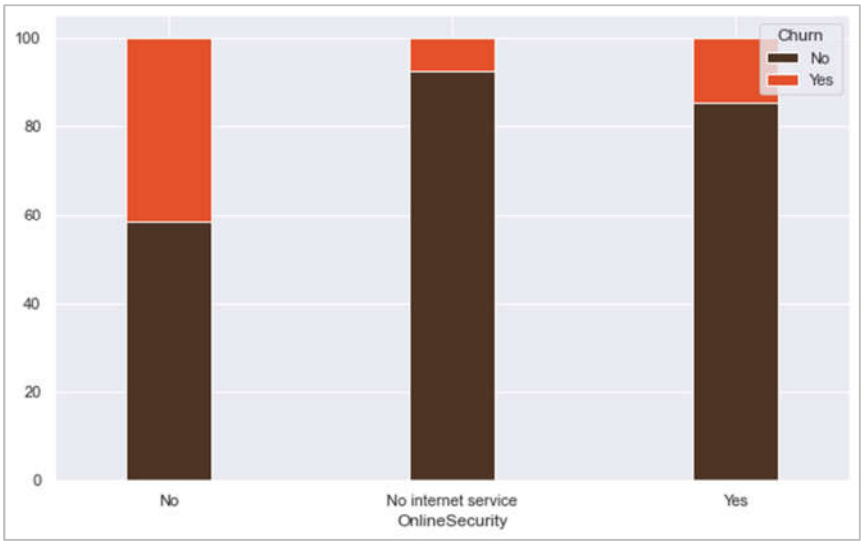
Study of effect of services on churn rate
Distribution of services



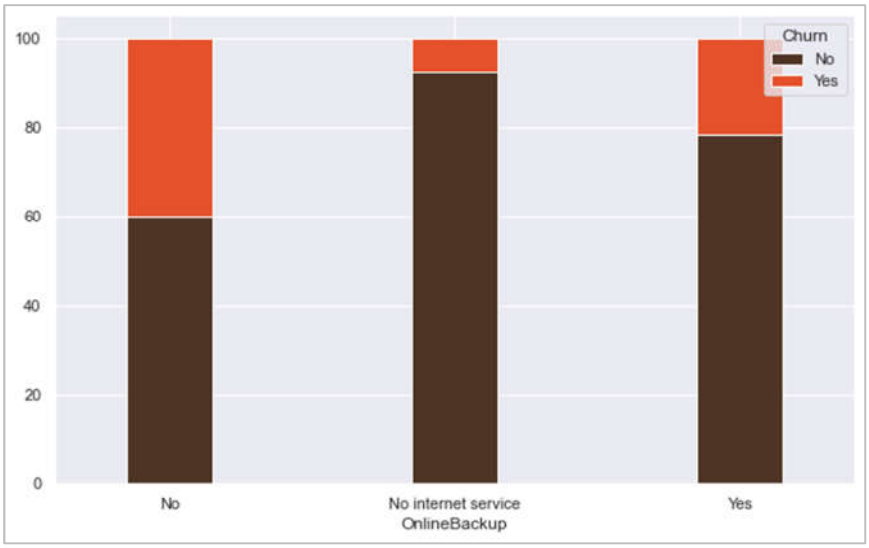
Relation of Fiber optics service with churn rate



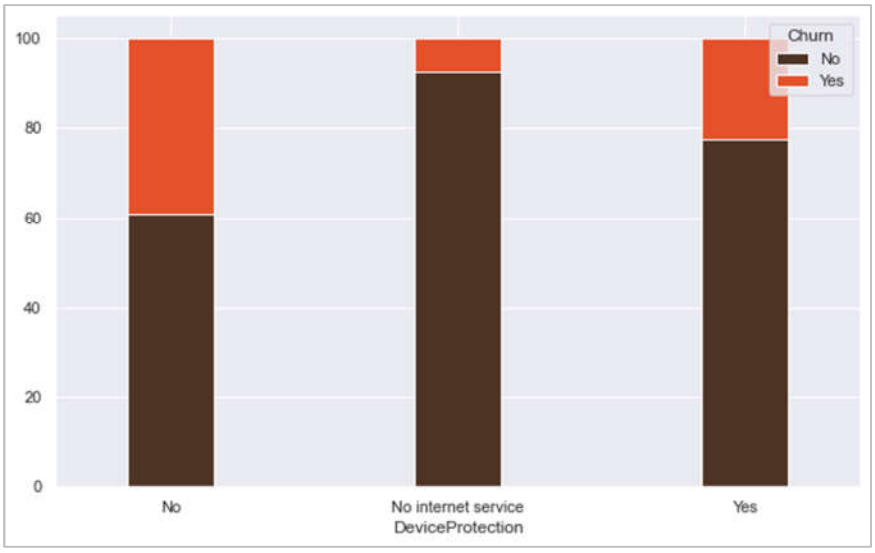
Relation of internet service subscription with churn rate



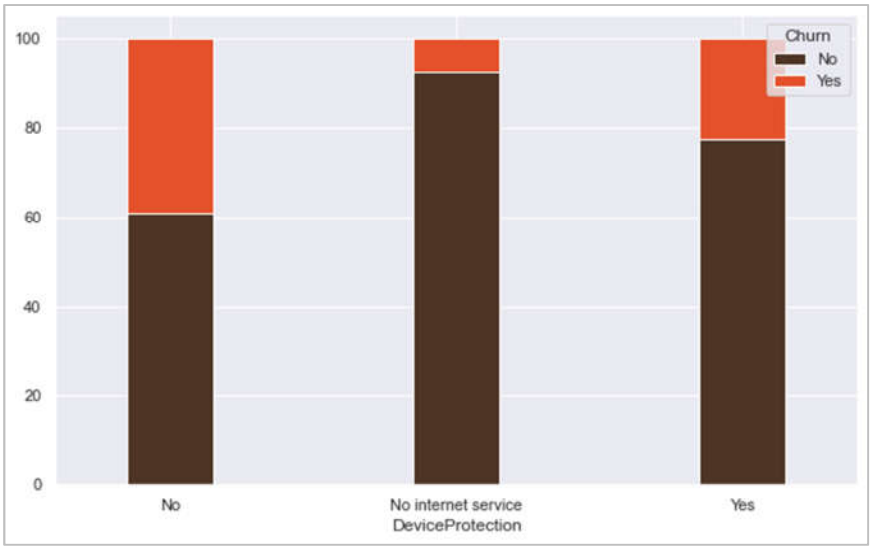
Relation of online backup service with churn rate



Relation of device protection service with churn rate



Relation of tech support service with churn rate



Solution

Though telecommunication companies have been putting in great efforts, they are still unable to conquer the churn battle because they aren't leveraging the bulk of data.

One of the reasons for this could be that most of the companies have been using traditional churn prediction models that have been heavily dependent only on the data gathered from transaction histories and demographics, but this method fails to integrate dynamic customer-generated input with real customer needs, wants, and wishes.

So, having a churn prediction model that can leverage customer interactions and is built on the latest technologies like machine learning and text analytics will help companies in analyzing the customer interactions at scale to understand customer behavior.

In this project the following analytical tools are used: -

- Market segmentation with **K-Means** clustering analysis
- Predicting customer churn with **Logistic Regression** and **Random Forest**

The Project we'll be dealing with:

- Exploring the Dataset
- Manipulation of data to Tenure Groups for later classification
- Use of Info provided to create Statistical Analysis of the state of Each Customer
- Understand the Cause of Customer Churning from the overall Subscribed Services & Other Info
- Visualizing the descriptive statistics of the whole Dataset
- Preprocess Data & Remove Un-Necessary Data
- Customer segmentation using clustering analysis
- Shuffle & Split Data
- Train & Test Both Random Forest & Logistic Regression Models
- Resample, Shuffle & Split Data then retrain
- Measure & Compare Final Scores and Improvements