

TELCO CHURN REPORT

**Analyzing Telco Customer Churn: Insights and Predictive Modelling**

**Introduction:**

Customer churn is a critical challenge for businesses, especially in the highly competitive telecommunications industry. This presentation report explores the ***Telco customer churn*** dataset and apply machine learning techniques to gain insights and build predictive models to identify factors that contribute to customer churn.

**Problem Statement**

Customer attrition is one of the biggest expenditures of any organization. Customer churn otherwise known as customer attrition or customer turnover is the percentage of customers that stopped using your company's product or service within a specified timeframe. For instance, if you began the year with 500 customers but later ended with 480 customers, the percentage of customers that left would be 4%. If we could figure out why a customer leaves and when they leave with reasonable accuracy, it would immensely help the organization to strategize their retention initiatives manifold. In this presentation we aim to find the likelihood of a customer leaving the organization, the key indicators of churn as well as the retention strategies that can be implemented to avert this problem.

**Understanding the Dataset:**

The Telco customer churn dataset contains information about customer demographics, services subscribed, contract details, and churn status. Our objective was to analyze this data and develop a model that predicts whether a customer is likely to churn or not.

Methodology

First of all, we examined the data structure including the variables, the presence of any missing values and any potential anomalies or outliers. This helped to inform our data preprocessing. We handled missing values in data set using deletion and imputation which made sure that our data is complete and clean hence ready for accurate model performance and training.

**Exploratory Data Analysis:**

* Next, we went ahead and started carrying out exploratory data analysis to understand the characteristics of the dataset. We examined the distribution of customer churn and found that approximately 1869 of the customers were of customers in the dataset churned, indicating a significant challenge for the Telco company. Also, Customers with shorter contract tenure churned more than those with longer tenure. The same way customers on monthly charges payment plan turn to churn more that those with yearly payment plan.
* Also, Among the internet service types, fiber optics dominate with 44%, with a high churn rate of 42%. DSL follows with 34%, with a churn rate of 16%. Customers without internet service make up 21%, with only a 10% churn rate. This suggests potential issues with fiber optic internet services leading to higher churn rates.
* We also found out that the "Month-to-month" contract type holds the largest share at 55%, with the highest churn rate at 44%. "One year" contracts represent 21%, with a churn rate of 10%. "Two year" contracts make up 24%, with only a 3% churn rate. This suggests higher turnover among new customers compared to loyal long-term ones.
* We further analyzed the data to identify key factors that influence customer churn. We investigated different demographic attributes, such as gender, age, and partner status, to understand their impact on churn. Additionally, we examined the relationship between churn and various services subscribed, contract types, and payment methods.

**Model Evaluation**

**Machine Learning Modelling:**

To predict customer churn, we employed several machine learning algorithms: Logistic Regression, Support Vector Machine and Random Forest Classifier. We also fixed class imbalances using resampling. We split the dataset into training and evaluation sets, trained each model on the training data, and evaluated their performance using the F1 score, a metric that balances precision and recall.

**Results and Insights:**

* Upon evaluating the models, we found that the Logistic Regression model achieved an overall accuracy of 0.7599, meaning it correctly predicts the churn status for approximately 75.99% of the customers in the test dataset. The accuracy score of the Random Forest model was approximately 0.896, indicating that it correctly predicts the class label for about 89.69% of the instances in the test dataset. The accuracy score of the SVM model is approximately 0.7531, indicating that it correctly predicts the class label for about 75.31%% of the instances in the test dataset.
* Random Forest Classifier turned out to perform better with an accuracy of 0.896 accuracy.

**Hyperparameter Tuning:**

To further enhance the performance of our best model, we performed hyperparameter tuning using GridSearchCV. By exploring different combinations of hyperparameters, we identified the best hyperparameters for the Logistic Regression model, SVM and random forests. This process improved the accuracy score, indicating better predictive performance.

* ***Best Hyperparameters for Logistic Regression: {'C': 0.1, 'penalty': 'l2'}***
* ***Best Hyperparameters for Random Forest: {'max\_depth': None, 'min\_samples\_leaf': 2,*** 
  + ***'min\_samples\_split': 2, 'n\_estimators': 200}***
* ***Best Hyperparameters for Support Vector Machine: {'C': 10, 'kernel': 'rbf'}***

Random Forest has the highest accuracy among the three models. With an accuracy of 0.87. This indicates that the model provides a reasonable balance between correctly identifying churned customers and minimizing false positives.

Also, we looked at the ROC of the three models. We can conclude that random forest performs the best in terms of ROC AUC compared to the other two models. It also outperforms the other models in ROC AUC Scores recording a highest ROC AUC of 96%.

The exists no difference in ROC AUC score for Logistic Regression and Support Vector Machine at 0.84 while Random Forest Classifier records the highest at 0.96, indicating that all three models perform reasonably well in distinguishing between the positive and negative classes.

**Conclusion:**

In conclusion, our analysis of the Telco customer churn dataset provided valuable insights into customer behavior and factors influencing churn. By leveraging machine learning techniques, we developed a predictive model that can identify potential churners and help the Telco company take proactive measures to retain customers. Understanding the impact of additional services and contract types can guide targeted marketing and service improvement strategies.

Based on the analysis of customer churn in the telecommunications industry, particularly with regards to Telco, several key insights have emerged:

* Tenure and Churn Rate: There's a clear inverse relationship between customer tenure and churn rate. Customers who have been with Telco for longer periods are less likely to churn. This suggests that as customers stay with the company over time, their satisfaction and loyalty tend to increase. It's indicative of a strong relationship between long-tenured customers and the company.
* Payment Method and Churn Rate: The choice of payment method appears to be correlated with churn rate. Customers who opt for electronic or mailed check payments exhibit higher churn rates compared to those using bank transfer or credit card with automatic payments. This implies that customers who utilize more automated and convenient payment methods may have a higher level of satisfaction and loyalty towards Telco.
* Paperless Billing and Churn Rate: Interestingly, customers who opt for paperless billing demonstrate a higher churn rate than those who prefer traditional billing methods. This suggests that there may be differing expectations or experiences among customers who embrace digital and paperless processes, influencing their decision to churn.
* These insights collectively highlight the importance of customer satisfaction, loyalty, and the customer experience in mitigating churn for Telco. Strategies aimed at enhancing customer satisfaction, particularly among newer customers, optimizing payment processes, and understanding the needs and preferences of digital-savvy customers can be crucial in reducing churn and fostering long-term relationships with customers.