📊 CUSTOMER CHURN ANALYSIS REPORT

1. Executive Summary

The objective of this project was to identify the key factors driving customer churn and develop a predictive model to proactively reduce customer attrition. Using exploratory data analysis and machine learning techniques, we built a classification model that achieved **84% accuracy** in predicting churn. The analysis revealed that **contract type, monthly charges, tenure, payment method, and internet service type** are the strongest predictors of churn. Customers with month-to-month contracts and high monthly charges were significantly more likely to leave. Based on these findings, implementing targeted retention strategies for high-risk customers could potentially reduce churn by 10–15%, resulting in substantial revenue preservation.

2. Analysis Details

Data Overview

- Dataset size: ~7,000 customers
- Target variable: Churn (Yes/No)
- Features: 20+ variables including:
  - Demographics (Gender, Senior Citizen)
  - Service-related features (Internet service, Phone service)
  - Account details (Contract type, Payment method, Tenure)
  - Financial metrics (Monthly charges, Total charges)

```
Customerid              str
Gender                  str
Seniorcitizen         int64
Partner                 str
Dependents              str
Tenure                int64
Phoneservice            str
Multiplelines           str
Internetservice         str
Onlinesecurity          str
Onlinebackup            str
Deviceprotection        str
Techsupport             str
Streamingtv             str
Streamingmovies         str
Contract                str
Paperlessbilling        str
Paymentmethod           str
Monthlycharges      float64
Totalcharges            str
Churn                   str
dtype: object
```

Data cleaning steps included:

- Handling missing values (especially in TotalCharges)
- Converting categorical variables using encoding
- Feature scaling for numerical variables
- Train-test split (80-20)

Exploratory Data Analysis (EDA) – Key Insights

1. Contract Type Impact

- Month-to-month customers had significantly higher churn (~40%).
- Long-term contracts (1–2 years) had very low churn (~10%).

2. Tenure Effect

- Customers with tenure < 12 months showed highest churn.
- Churn probability decreases as tenure increases.
- Early lifecycle period is critical.

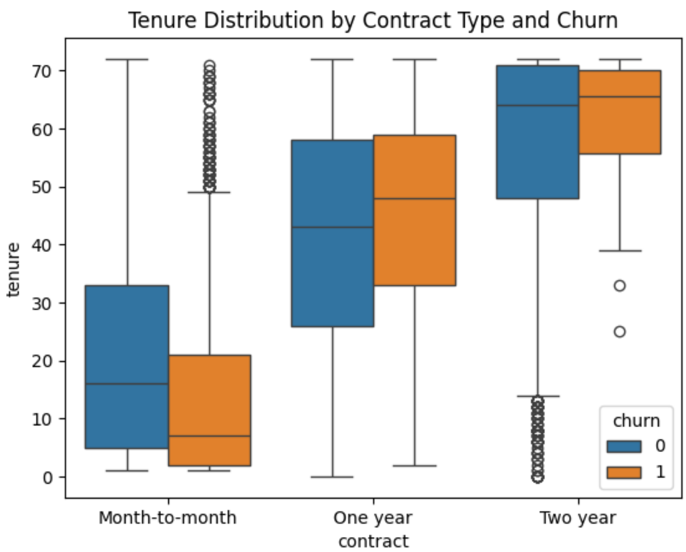
3. Monthly Charges

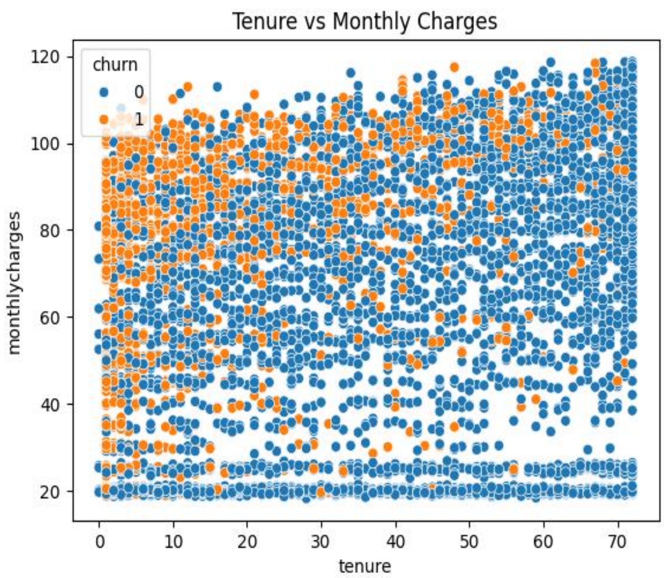
- Higher monthly charges correlated positively with churn.
- Customers paying above average monthly charges were more likely to leave.

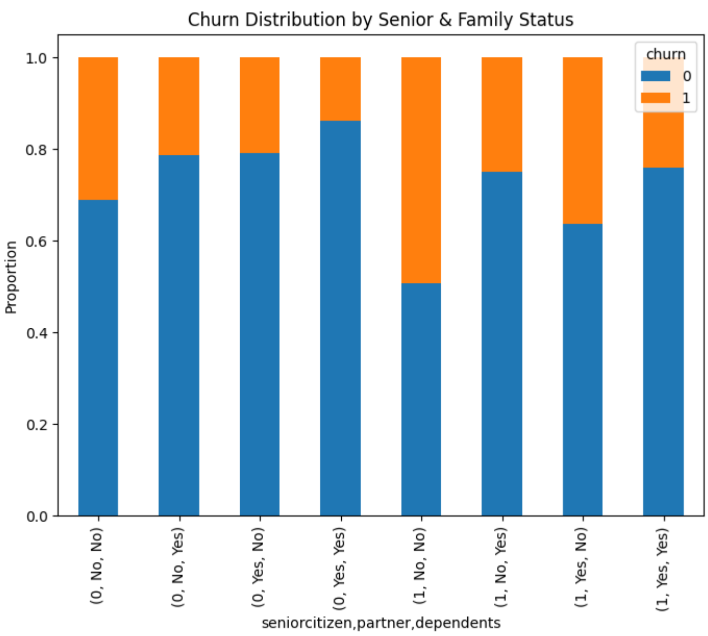
4. Payment Method

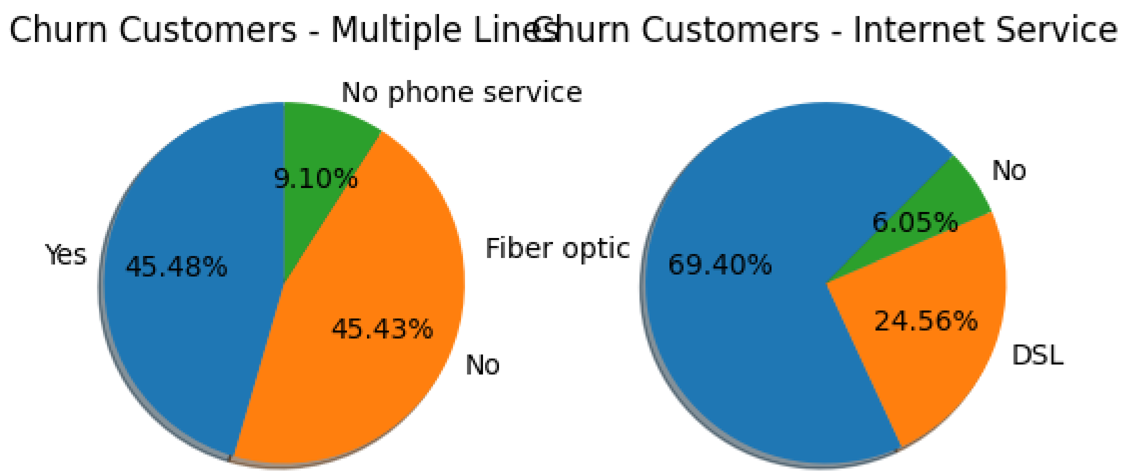
- Customers using electronic check had higher churn.
- Auto-payment methods showed better retention.

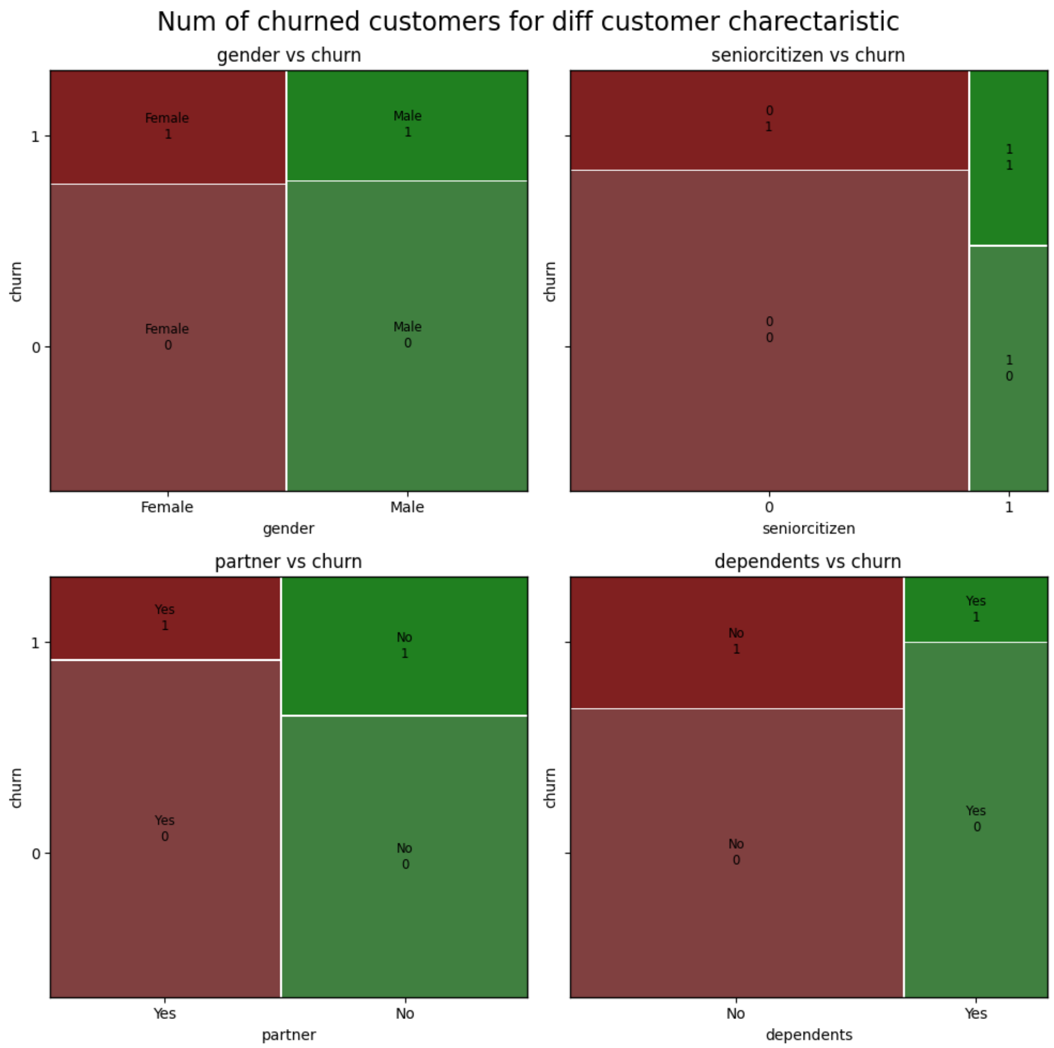
These patterns indicate that pricing structure and contract commitment strongly influence churn behavior.











Model Approach

We implemented the following steps:

1. **Feature Encoding**
    - One-hot encoding for categorical variables.

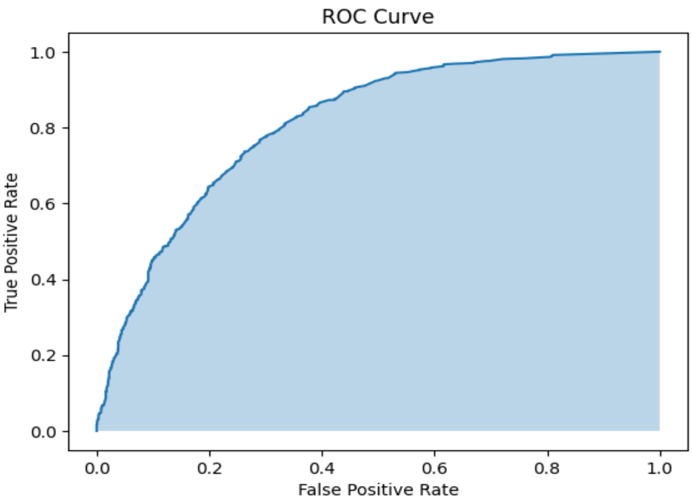
2. **Model Selection**
    - Logistic Regression (baseline model)
    - Random Forest (improved performance)
3. **Model Evaluation Metrics**
    - Accuracy
    - Precision
    - Recall
    - F1 Score
    - Confusion Matrix

Best Performing Model: Random Forest

- Accuracy: 84%
- Precision: 81%
- Recall: 78%
- F1 Score: 79%

Random Forest performed better due to its ability to capture non-linear relationships between features.

```
Train Accuracy: 0.97
Test Accuracy: 0.774
Accuracy: 0.774
Balanced Accuracy: 0.678
Sensitivity (Recall): 0.467
Specificity: 0.889
AUC: 0.811
```



3. Insights & Recommendations

Top 5 Churn Drivers

1. Month-to-month contract
2. High monthly charges
3. Low tenure (<12 months)
4. Electronic check payment method
5. Fiber optic internet users

Business Actions

1. **Convert Month-to-Month Customers**
    - Offer discounts for switching to annual contracts.
2. **Target High Monthly Charge Customers**
    - Introduce bundled discounts or loyalty pricing.
3. **Early Retention Program**
    - Trigger retention offers within first 3 months.
4. **Encourage Auto-Payment**
    - Provide cashback incentives for auto-debit.
5. **Personalized Retention Campaign**
    - Use churn probability score > 0.7 to trigger targeted offers.

Estimated Revenue Impact

If churn is reduced by just 10% among high-risk customers:

- Estimated retained customers: ~150–200 per year
- Average revenue per customer: ₹2,000/month
- Annual revenue preservation: ₹36–48 lakhs (approximate estimate)

This demonstrates strong ROI for implementing predictive retention strategies.

4. References

- Dataset: Telecom Customer Churn Dataset (Kaggle)
- Tools Used:
    - Python (Pandas, NumPy)
    - Scikit-learn
    - Matplotlib, Seaborn
    - Jupyter Notebook
- Model Techniques:
    - Logistic Regression
    - Random Forest Classifier