PREDICTING CUSTOMER CHURN FOR PERSONALIZED RETENTION STRATEGIES

Vema Dondeti

Predictive Analytics

INTRODUCTION – UNDERSTANDING CUSTOMER CHURN

- **Definition:** Customer churn occurs when customers discontinue their relationship with a company, either by canceling a service, closing an account, or switching to a competitor.
- **Why it matters:** Acquiring new customers is significantly more expensive than retaining existing ones. Reducing churn boosts loyalty and profits.
- **Challenges:**
  - Hard to predict due to complex behaviors.
  - Delayed feedback makes it tricky to act fast.
  - Without data, retention efforts can be costly and ineffective.

PROJECT OBJECTIVE

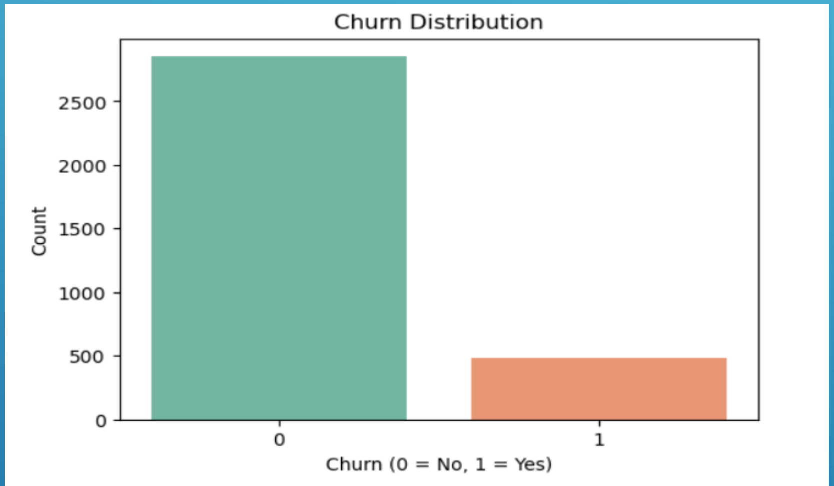
- **Objective:** Develop a predictive model to identify customers at risk of churn.
- **Key Goals:**
  - Understand key factors driving churn.
  - Improve customer retention through data-driven insights.
  - Provide actionable recommendations for personalized retention strategies.

DATASET OVERVIEW

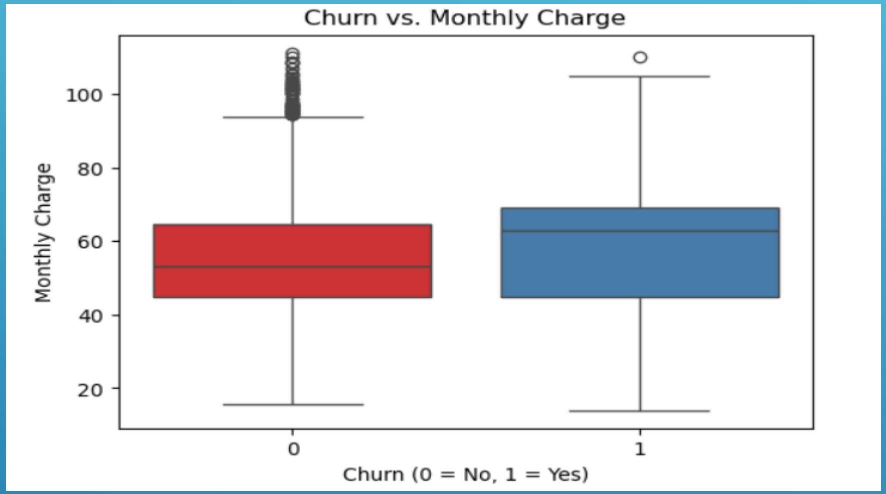
- **Source:** Telecom Customer Churn dataset from Kaggle
- **Size:** 3,333 customer records, 11 columns
- **Key Features:**
  - Account Duration: AccountWeeks (length of customer relationship)
  - Service Usage: DataPlan, DataUsage, RoamMins (roaming minutes), DayMins (call duration)
  - Billing & Charges: MonthlyCharge, OverageFee
  - Customer Support: CustServCalls (support calls), ContractRenewal
- **Target Variable:** Churn (1 = Customer Left, 0 = Customer Retained)

DATA VISUALIZATIONS - CHURN TRENDS

Churn Distribution



Churn vs. Monthly Charge

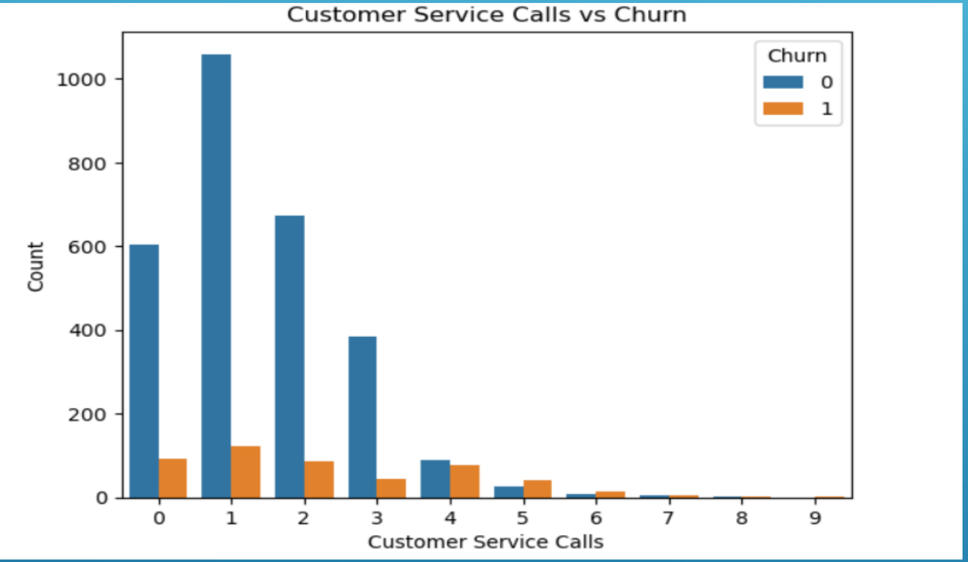


Key Insight:

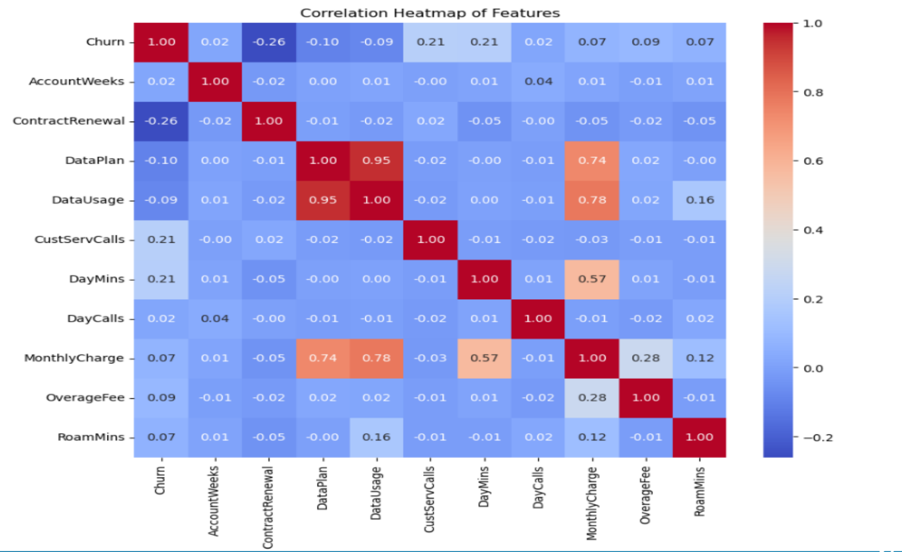
- More customers stayed (0) than left (1), indicating class imbalance.
- Higher monthly charges may contribute to churn, but other factors play a role.

DATA VISUALIZATIONS - CUSTOMER BEHAVIOR

Customer Service Calls vs. Churn



Heatmap of Feature Correlations



Key Insight:

- Frequent customer service calls are linked to higher churn.
- Correlation heatmap reveals key churn predictors.

DATA PREPARATION

- Dropped Features: Removed ‘DataPlan’ due to multicollinearity.
- Handling Missing Values: Imputed numerical columns with the median.
- Feature Engineering: Created new features based on customer behavior.
- Outlier Handling: Capped extreme values at the 95th percentile.
- Standardization: Applied StandardScaler for consistency.
- Class Imbalance: Used SMOTE to balance churned vs. retained customers.

MODEL SELECTION & TRAINING

- Models Tested:
  - Logistic Regression - Baseline model for interpretability
  - Random Forest - Reduces overfitting by combining multiple decision trees.
  - Extreme Gradient Boosting (XGBoost) - Optimizes classification by minimizing errors iteratively.
- Training Approach:
  - Data split: 80% training, 20% testing.
  - Models trained on SMOTE-balanced dataset to handle class imbalance.

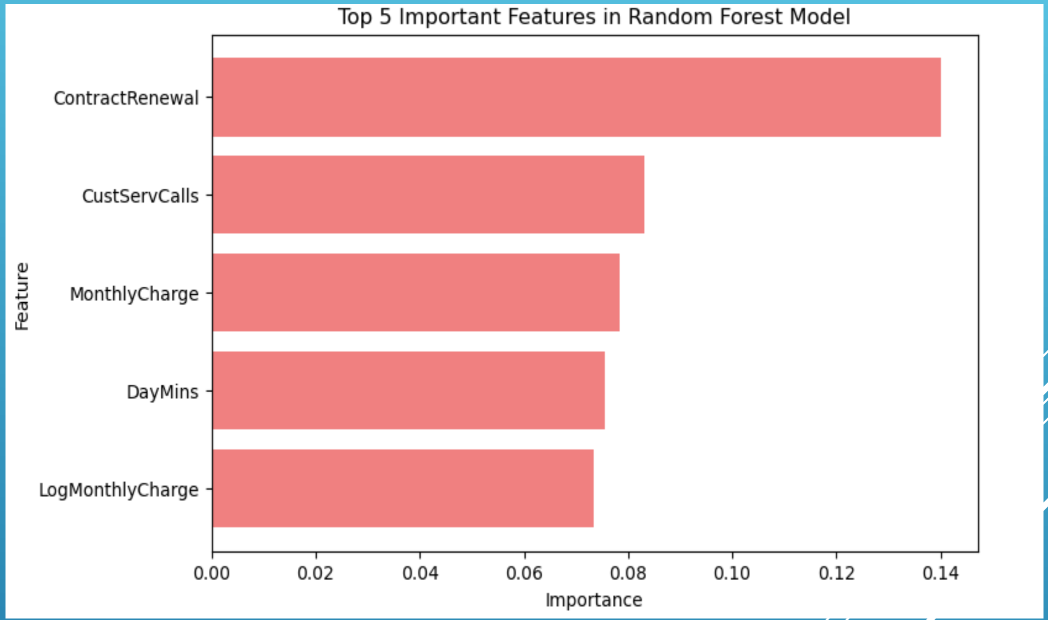
MODEL PERFORMANCE & SELECTION

Model	Precision	Recall	AUC-ROC	Notes
Random Forest	66%	**72%**	**0.83**	Best balance, selected for deployment
Logistic Regression	45%	81%	0.78	High recall but too many false positives
XGBoost	71%	67%	0.81	High precision but missed more churners

Final Choice: Random Forest, as it effectively identifies churners while minimizing false positives

RECOMMENDATIONS FOR REDUCING CHURN

- Encourage Contract Renewals –
  - Offer discounts or perks.
- Improve Customer Support –
  - Resolve complaints efficiently.
- Optimize Pricing & Plans –
  - Recommend better plans based on usage.
- Monitor Usage Behavior –
  - Identify at-risk customers early.



CONCLUSION & NEXT STEPS

- Key Takeaways:
  - Predicting churn enables businesses to implement targeted retention strategies and reduce customer churn.
  - Random Forest delivered the best balance of precision and recall, making it the most effective for identifying at-risk customers.
  - Improving contract renewals, enhancing customer support, and optimizing pricing can significantly lower churn rates.
- Next Steps:
  - Enhance model performance with additional behavioral and demographic features, and test real-world interventions.

REFERENCES

- Prabadevi, B., Shalini, R., & Kavitha, B. (2023). Customer churning analysis using machine learning algorithms. *International Journal of Intelligent Networks*, *4*, 145–154. https://doi.org/10.1016/j.ijin.2023.05.005
- Ouko, A. (2024, December 10). Customer churn prediction using Machine Learning - Allan Ouko - Medium. *Medium*. https://medium.com/@allanouko17/customer-churn-prediction-using-machine-learning-ddf4cd7c9fd4
- *Customer churn*. (2020, March 23). Kaggle. https://www.kaggle.com/datasets/barun2104/telecom-churn/data
- https://www.ibm.com/think/topics/customer-churn