TELECOM CUSTOMER CHURN PREDICTION

PHASE 3 PROJECT

BY:TYRESE MUGO

PROJECT OVERVIEW

- Predict which telecom customers are likely to churn- Identify customers who are at risk of leaving the service.
- Reduce revenue loss and improve retention- Retaining customers is more cost-effective than acquiring new ones.
- Help marketing teams target high-risk customers- Allows personalized campaigns to keep valuable customers.

BUSINESS PROBLEM

01 Customer churn reduces revenue and increases costs

02 Retaining existing customers is cheaper than acquiring new ones

03 Identify high-risk customers early

STAKEHOLDER

- Primary stakeholder: Telecom company management and marketing teams
- They need actionable insights to make decisions that reduce churn.
- Decisions supported by model:
- Identify customers to target for retention campaigns and personalized offers.
- Focus: Use predictive insights to improve customer satisfaction and loyalty

BUSINESS GOALS & IMPACT

Increase customer retention rates
Keeping existing customers improves long-term revenue stability.

Reduce revenue loss due to churn
Proactive retention lowers the financial impact of lost customers.

Reduce revenue loss due to churn
Proactive retention lowers the financial impact of lost customers.

DATASET OVERVIEW

SOURCE: KAGGLE TELECOM CHURN DATASET
CONTAINS REAL CUSTOMER RECORDS FROM A TELECOM COMPANY.
RECORDS: 3,333 CUSTOMERS
PROVIDES ENOUGH DATA FOR TRAINING AND TESTING MODELS.
FEATURES: 21 FEATURES INCLUDING USAGE METRICS, SERVICE PLANS, ACCOUNT INFO, AND CUSTOMER SERVICE INTERACTIONS
TARGET VARIABLE: CHURN (YES/NO)
INDICATES WHETHER A CUSTOMER HAS LEFT THE SERVICE.

EXPLORATORY DATA ANALYSIS (EDA) INSIGHTS

- CLASS IMBALANCE: FEWER CHURNED CUSTOMERS THAN RETAINED
- THIS IMPACTS EVALUATION METRICS — ACCURACY ALONE IS INSUFFICIENT.
- HIGH CUSTOMER SERVICE CALLS → HIGHER CHURN
- CUSTOMERS WHO CONTACT SUPPORT FREQUENTLY ARE MORE LIKELY TO LEAVE.
- INTERNATIONAL PLAN → HIGHER CHURN LIKELIHOOD
- CERTAIN PLANS MAY BE ASSOCIATED WITH DISSATISFACTION OR HIGHER COSTS.
- PURPOSE: IDENTIFY PATTERNS AND IMPORTANT FEATURES FOR MODELING
- HELPS INFORM FEATURE SELECTION AND MODEL CHOICE.

DATA PREPARATION

CATEGORICAL VARIABLES ONE-HOT ENCODED
CONVERTS CATEGORIES INTO NUMERIC FORM SUITABLE FOR MACHINE LEARNING.
TARGET SEPARATED FROM FEATURES
PREVENTS DATA LEAKAGE DURING TRAINING.
STRATIFIED TRAIN-TEST SPLIT (75%-25%)
ENSURES THE SAME CHURN PROPORTION IN TRAINING AND TESTING SETS.

MODELING APPROACH

- BASELINE MODEL: LOGISTIC REGRESSION
- SIMPLE AND INTERPRETABLE, PROVIDES A REFERENCE PERFORMANCE.
- NON-PARAMETRIC MODEL: DECISION TREE
- CAPTURES NON-LINEAR RELATIONSHIPS AND FEATURE INTERACTIONS.
- ITERATIVE APPROACH
- COMPARE BASELINE AND ALTERNATIVE MODELS TO SELECT THE BEST PERFORMING ONE.

MODEL EVALUATION

- KEY METRIC: RECALL
- FOCUSED ON CORRECTLY IDENTIFYING CUSTOMERS LIKELY TO CHURN.
- DECISION TREE: HIGHER RECALL FOR CHURNED CUSTOMERS
- MORE EFFECTIVE FOR PROACTIVE RETENTION CAMPAIGNS.
- LOGISTIC REGRESSION: SLIGHTLY LOWER RECALL BUT EASIER TO INTERPRET
- USEFUL FOR EXPLAINING MODEL BEHAVIOR TO BUSINESS STAKEHOLDERS.

RECOMMENDATIONS & NEXT STEPS

- FLAG HIGH-RISK CUSTOMERS USING THE DECISION TREE MODEL
- PRIORITIZE THEM FOR RETENTION CAMPAIGNS AND PERSONALIZED OFFERS.
- FOCUS ON CUSTOMERS WITH:
  - HIGH SERVICE USAGE
  - INTERNATIONAL PLAN
  - FREQUENT CUSTOMER SERVICE CALLS
- FUTURE IMPROVEMENTS:
  - USE ENSEMBLE MODELS FOR BETTER ACCURACY
  - INCLUDE DEMOGRAPHIC AND CONTRACT DATA
  - TRACK LONGITUDINAL BEHAVIOR FOR IMPROVED PREDICTIONS

THANK YOU!!!