

AI-DRIVEN CUSTOMER CHURN RISK & RETENTION DECISION SYSTEM

DATABRICKS LAKEHOUSE (COMMUNITY EDITION)

PROBLEM STATEMENT

CUSTOMER CHURN DIRECTLY IMPACTS REVENUE

REACTIVE RETENTION IS EXPENSIVE

BUSINESSES NEED EARLY, EXPLAINABLE RISK SIGNALS

OBJECTIVE

IDENTIFY CHURN RISK EARLY AND ENABLE RETENTION DECISIONS

FOCUS ON DECISION INTELLIGENCE, NOT JUST PREDICTION.

DATASET USED

TELECOM CUSTOMER CHURN DATASET (KAGGLE)

KEY FEATURES

Customer demographics

Services & contracts

Billing information

Churn indicator (Yes / No)

Public dataset, real-world structure

DATABRICKS LAKEHOUSE DESIGN

BRONZE → SILVER → GOLD → DECISION LAYER

BRONZE: RAW INGESTION

SILVER: CLEANED & STANDARDIZED DATA

GOLD: ANALYTICS + AI-READY DATA

DECISION LAYER: CHURN RISK SCORING

BRONZE LAYER

- **RAW CUSTOMER DATA STORED AS DELTA TABLE**
- **NO TRANSFORMATIONS APPLIED**
- **ENSURES TRACEABILITY AND AUDITABILITY**

SILVER LAYER

- **HANDLED MISSING AND MALFORMED VALUES**
- **FIXED DATA TYPES AND STANDARDIZED CATEGORIES**
- **PREPARED TRUSTED DATA FOR ANALYTICS**

GOLD LAYER

- **RULE-BASED CHURN RISK SCORING**

- **RISK SIGNALS INCLUDE:**

- **TENURE**
- **CONTRACT TYPE**
- **MONTHLY CHARGES**
- **INTERNET SERVICE**
- **PAYMENT METHOD**

DESIGN CHOICE: DECISION-FIRST APPROACH

- **FOCUSED ON EXPLAINABLE CHURN RISK SCORING, NOT A BLACK-BOX MODEL**
- **PRIORITIZED BUSINESS TRUST AND INTERPRETABILITY**
- **MIRRORS REAL-WORLD SYSTEMS WHERE RULES + DATA COME BEFORE ML**
- **ARCHITECTURE REMAINS ML-READY FOR FUTURE EXTENSION**

CHURN RISK SCORING LOGIC

- **CUSTOMER TENURE**
- **CONTRACT TYPE (MONTH-TO-MONTH VS LONG-TERM)**
- **INTERNET SERVICE TYPE**
- **PAYMENT METHOD**
- **MONTHLY BILLING AMOUNT**

FROM RISK SCORES TO BUSINESS ACTIONS

- **LOW RISK → NO IMMEDIATE ACTION**
- **MEDIUM RISK → ENGAGEMENT & MONITORING**
- **HIGH RISK → RETENTION OFFERS / INTERVENTION**

FEATURE SELECTION STRATEGY (ML EXTENSION)

- ORIGINAL DATASET CONTAINED MANY CATEGORICAL VARIABLES
- ONE-HOT ENCODING CAUSED FEATURE SPACE EXPLOSION
- FEATURE SELECTION APPLIED TO ENSURE ML STABILITY

DROPPED COLUMNS:

ONLINESECURITY, ONLINEBACKUP
DEVICEPROTECTION, TECHSUPPORT
STREAMINGTV, STREAMINGMOVIES
MULTIPLELINES, PHONESERVICE
PARTNER, DEPENDENTS

RETAINED FEATURES:

TENURE, MONTHLYCHARGES,
TOTALCHARGES
SENIORCITIZEN
CONTRACT, INTERNETSERVICE
PAYMENTMETHOD, GENDER

KEY BUSINESS INSIGHTS

- **HIGH-RISK CUSTOMERS TEND TO HAVE SHORTER TENURE**
- **MONTH-TO-MONTH CONTRACTS SHOW HIGHER CHURN RISK**
- **HIGHER MONTHLY CHARGES CORRELATE WITH CHURN**
- **PAYMENT METHOD INFLUENCES CUSTOMER STABILITY**

WHAT THIS PROJECT DEMONSTRATES

- **DATABRICKS LAKEHOUSE ARCHITECTURE**
- **DATA CLEANING & TRANSFORMATION**
- **SQL-BASED DECISION LOGIC**
- **EXPLAINABLE AI THINKING**
- **BUSINESS-ORIENTED DATA ENGINEERING**

LIMITATIONS & TRADE-OFFS

- **RULE-BASED SYSTEM MAY MISS COMPLEX PATTERNS**
- **NO AUTOMATED LEARNING FROM NEW DATA**
- **CHOSEN INTENTIONALLY FOR EXPLAINABILITY AND TRUST**

FUTURE ENHANCEMENTS

- **REPLACE RULES WITH ML CHURN PREDICTION MODEL**
- **USE SPARK ML / MLFLOW**
- **ADD GENERATIVE AI FOR CHURN EXPLANATION SUMMARIES**
- **AUTOMATE RETENTION RECOMMENDATIONS**

THANK YOU

PRESENTED BY - LIZA DEKA