

TELECOM CUSTOMER CHURN ANALYSIS (POWER BI + SQL + MACHINE LEARNING)

PROJECT GOALS

Create an entire ETL (Extract Transform Load) process in a database & a Power BI dashboard to utilize the customer data and achieve goals:

1. Analyze Customer Data at below levels
 - a) Demographic
 - b) Geographic
 - c) Payment & Account Information
 - d) Services
2. Study Churner Profile & Identify Areas for Implementing Marketing Campaigns
3. Identify a Method to Predict Future Churners

METRICS REQUIRED

1. Total Customers
2. Total Churn & Churn Rate
3. New Joiners

ETL (EXTRACT TRANSFORM LOAD) FRAMEWORK

Our framework uses below components:

- CSV file – This is our source file
- SQL Server Management Studio – we will use its inbuilt **Import wizard** to **transform & load the data**
- SQL Server Database – This is where our final data will be **loaded & host** our data warehouse, tables & views for final usage

Power BI Visualization

Summary Page

1. Top Card
 - a. Total Customers
 - b. New Joiners
 - c. Total Churn
 - d. Churn Rate%

2. Demographic

- a. Gender – Churn Rate
- b. Age Group – Total Customer & Churn Rate

3. Account Info

- a. Payment Method – Churn Rate
- b. Contract – Churn Rate
- c. Tenure Group – Total Customer & Churn Rate

4. Geographic

- a. Top 5 State – Churn Rate

5. Churn Distribution

- a. Churn Category – Total Churn
- b. Tooltip : Churn Reason – Total Churn

6. Service Used

- a. Internet Type – Churn Rate
- b. prod_Service >> Services – Status – % RT Sum of Churn Status

Churn Reason Page (Tooltip)

- 1. Churn Reason – Total Churn

Predict Customer Churn

For predicting customer churn, we will be using a widely used Machine Learning algorithm called RANDOM FOREST.

MACHINE LEARNING

Machine learning is a branch of artificial intelligence (AI) that enables computers to learn from and make decisions based on data without being explicitly programmed for specific tasks. It involves using algorithms and statistical models to find patterns in data and improve performance over time as find patterns in data and improve performance over time as more data is processed.

PREDICTED ANALYTICS

Predictive analytics is a type of data analysis that uses statistical algorithms, machine learning techniques, and historical data to make predictions about future events. It identifies patterns in the data to forecast outcomes and trends, helping organisations make data-driven decisions.

RANDOM FOREST ALGORITHM (Most Widely Used Model When It Comes to Churn Analysis)

A random forest is a machine learning algorithm that consists of multiple decision trees. Each decision tree is trained on a random subset of the data and features. The final prediction is made by averaging the predictions (in regression tasks) or taking the majority vote (in classification tasks) from all the trees in the forest. This ensemble approach improves the accuracy and robustness of the model by reducing the risk of overfitting compared to using a single decision tree.

Data Preparation for ML model

Let us first import views in an Excel file.

- o Go to Data >> Get Data >> SQL Server Database
- o Enter the Server Name & Database name to connect to SQL Server
- o Import both vw_ChurnData & vw_JoinData
- o Save the file as Churn Prediction & Joined Prediction

Power BI Visualization of Predicted Data

Import CSV Data or Load Predicted data in SQL server & connect to server

Churn Prediction Page (Using New Predicted Data)

1. Right Side Grid
 - a. Customer ID
 - b. Monthly Charge
 - c. Total Revenue
 - d. Total Refunds
 - e. Number of Referrals
2. Demographic
 - a. Gender – Churn Count
 - b. Age Group – Churn Count
 - c. Marital Status – Churn Count
3. Account Info
 - a. Payment Method – Churn Count
 - b. Contract – Churn Count
 - c. Tenure Group – Churn Count
4. Geographic
 - a. State – Churn Count