Phase - 1

Defining the problem statements

| DATE	28 - 09 - 2023
TEAM ID	8939
PROJECT NAME	8301 – CUSTOMER CHURN PREDICTION
TEAM NAME	Proj_207142_Team_1

CUSTOMER CHURN PREDICTION

- DATA ANALYTICS

Problem Statement :

Develop a customer churn prediction model for XYZ Telecom to identify and proactively retain customers at risk of leaving the service. By analyzing historical customer data, behavioral patterns, and churn labels, the objective is to reduce churn rates, enhance customer satisfaction, and optimize customer retention strategies. The scope includes data collection and preprocessing, model selection, and evaluation using key metrics like accuracy, precision, and recall. Success criteria involve achieving a significant reduction in churn rates and improving the model's accuracy in classifying churners and non-churners. The actionable insights derived from the model will guide the implementation of targeted retention campaigns and personalized offers. Continuous monitoring and model adaptation will be integral to the project's long-term success.

Customer churn prediction is an important part of data analytics that uses data and predictive modeling to identify which customers are likely to end their relationship with a business or unsubscribe from a service. The goal is to take steps to reduce churn, as it is often more cost-effective to retain existing customers than to acquire new ones.

7 Steps to Predicting Customer Churn

1. Data Preparation
2. Feature Selection
3. Model Selection
4. Data Splitting
5. Model Training
6. Churn Prediction
7. Threshold Selection
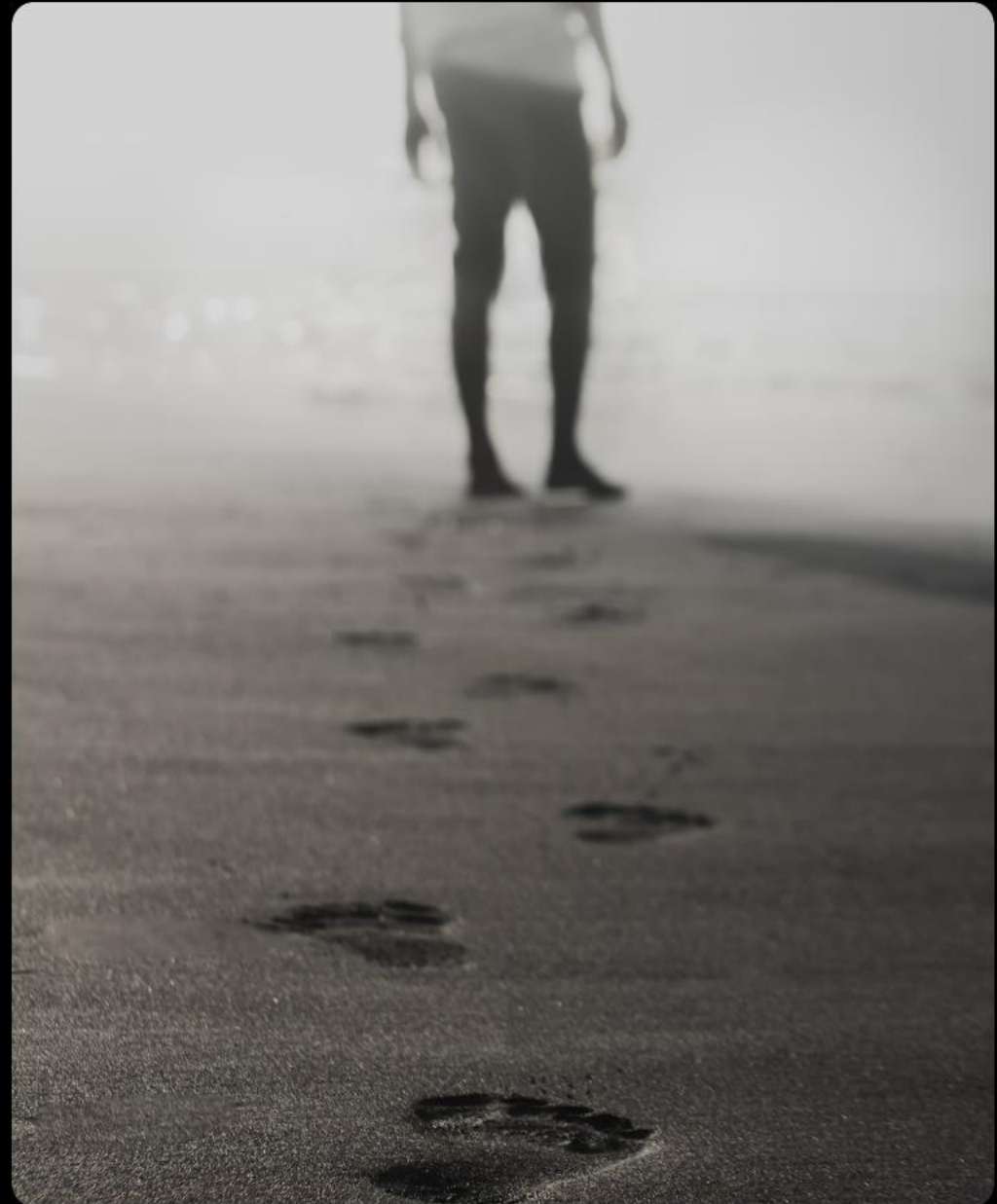
Data Preparation

Data collection and preparation involves gathering relevant data on customer behavior, interactions, and demographics, and preprocessing it to check for quality and suitability for analysis. This may include transaction history, customer support interactions, website/app usage, survey responses, and more, as well as dealing with missing values, outlier detection, and feature engineering.

Feature Selection

2. **Feature Selection and Engineering**:

- Identify relevant features: Select the most important variables that can potentially influence churn. These may include customer demographics, purchase history, usage patterns, and engagement metrics.

- Feature engineering: Create new features or transform existing ones to extract meaningful insights. For example, calculating customer lifetime value (CLV) or churn risk scores.

Model Selection

3. **Model Selection**:

- Choose appropriate machine learning algorithms: Common choices include logistic regression, decision trees, random forests, support vector machines, and neural networks.

- Evaluate model performance: Use metrics like accuracy, precision, recall, F1-score, and area under the ROC curve (AUC-ROC) to assess the model's effectiveness.

Data Splitting

4. **Data Splitting**:

- Split the dataset into training, validation, and testing sets to train the model, tune hyperparameters, and evaluate its generalization performance.

Model Training

5. **Model Training and Validation**:

- Train the chosen model on the training data.

- Use the validation set to fine-tune hyperparameters and avoid overfitting.

Churn Prediction

6. **Churn Prediction**:

- Use the trained model to predict churn probabilities for each customer in the testing or real-world data.

Threshold Selection

7. **Threshold Selection**:

- Set a probability threshold for classifying customers as likely churners or non-churners. This threshold can be adjusted to balance false positives and false negatives based on business goals.

Actionable Insights

Companies can use churn prediction strategies to reduce customer acquisition costs and improve their bottom line. This involves translating model predictions into actionable insights, monitoring and updating models, and creating a feedback loop between data analytics and business operations.

CONCLUSION :

Customer churn prediction, a key facet of data analytics, empowers businesses to foresee and mitigate customer attrition. It relies on data collection, preprocessing, and the selection of appropriate machine learning models for accurate predictions. By identifying customers likely to leave, companies can proactively tailor retention strategies, such as personalized offers, ultimately improving customer loyalty and reducing acquisition costs. Continuous model monitoring and iteration are crucial to adapt to changing customer behavior. Establishing a feedback loop between analytics and operations ensures that predictive insights translate into effective actions, bolstering a company's bottom line and fostering growth.