**Churn Prediction using Machine Learning**

**Abstract:**

Customer churn is a critical challenge for businesses, particularly in sectors such as telecommunications, banking, and subscription-based services. Churn refers to customers discontinuing their engagement with a service, leading to revenue loss and increased customer acquisition costs. This project presents a machine learning-based churn prediction model that analyzes historical customer data, including demographic details, service usage patterns, transaction history, and customer support interactions, to identify potential churners.

A dataset is preprocessed by handling missing values, feature scaling, and encoding categorical variables before applying various classification algorithms such as Logistic Regression, Decision Trees, Random Forest, Support Vector Machines (SVM), and Gradient Boosting techniques (XGBoost, LightGBM). Feature selection techniques are employed to identify the most influential attributes contributing to customer churn. The model is evaluated using key performance metrics such as accuracy, precision, recall, F1-score, and AUC-ROC curve to determine its effectiveness.

Additionally, insights derived from model explainability techniques such as SHAP (Shapley Additive Explanations) and feature importance ranking help businesses understand the primary reasons for churn and implement targeted retention strategies. The final model is integrated into a user-friendly interface that enables business decision-makers to predict customer churn and take proactive measures, such as personalized offers, improved customer service, or loyalty programs, to retain valuable customers.

This system provides a data-driven approach to churn prediction, allowing companies to make informed decisions and enhance customer satisfaction, ultimately improving long-term business sustainability.