

Customer Churn Prediction Project Documentation

Introduction

Welcome to the documentation for the Customer Churn Prediction project. This document provides an overview of the project's objectives, methodology, key findings, and results. The goal of this project was to develop a machine learning model that predicts customer churn based on historical customer data.

Project Objective

The primary objective of this project was to build a predictive model capable of identifying customers who are likely to churn. By leveraging historical customer data, we aimed to contribute to Sunbase's efforts in understanding customer behaviour and ensuring customer satisfaction.

Methodology

The project followed a structured machine learning project pipeline, encompassing the following key steps:

1. Data Preprocessing: Loaded the provided dataset, handled missing data, addressed outliers, encoded categorical variables, and split the data into training and testing sets.
2. Feature Engineering: Generated relevant features from the dataset to enhance model prediction accuracy. Applied feature scaling and normalization where necessary.
3. Model Building: Selected suitable machine learning algorithms, trained and validated the models, and evaluated performance using accuracy, precision, recall, and F1-score.
4. Model Optimization: Fine-tuned model hyperparameters using techniques such as grid search and cross-validation to enhance predictive performance.
5. Model Deployment: Deployed the trained model in a simulated production environment, demonstrating its capability to make predictions on new customer data.

Code Implementation

The project's code was implemented using Python and relevant machine learning libraries, including scikit-learn for model development, pandas for data manipulation, and matplotlib/seaborn for visualizations. The code was organized into a Jupyter Notebook, allowing for an interactive and iterative approach to development.

Project Report

A detailed project report was prepared to document the approach, decisions, and findings. The report includes the following sections:

1. Introduction: Brief overview of the project's context and objectives.
2. Data Preprocessing: Explanation of how missing data, outliers, and categorical variables were handled.

3. Feature Engineering: Details on the features created and their significance.
4. Model Selection and Building: Rationale for selecting specific algorithms, model training and validation, and performance evaluation.
5. Model Optimization: Description of hyperparameter tuning techniques and their impact on model performance.
6. Model Deployment: Overview of the simulated deployment environment and the model's performance in it.
7. Results and Insights: Presentation of model performance metrics, visualizations, and key takeaways.

Deliverables

The project's deliverables include:

1. Code Implementation: Jupyter Notebook containing the complete code for the project.
2. Project Report: A comprehensive report detailing the project's approach, findings, and insights.
3. Performance Metrics and Visualizations: Included in the project report to showcase model performance.

Conclusion

This project successfully developed a machine learning model capable of predicting customer churn based on historical customer data. The structured pipeline, thorough data preprocessing, insightful feature engineering, effective model building, and rigorous optimization contributed to achieving this objective.

For inquiries or further information, please contact [Sejal Mankar] at [sejalmankar444@gmail.com].

Thank you for your interest in the Customer Churn Prediction project.