**Name: -** Vatsal Rameshbhai Gohel

**Mail id: -** [**gohel.vatsal11@gmail.com**](mailto:gohel.vatsal11@gmail.com)

**Contact Detail: -** +1-774-253-3960

**LinkedIn Id: -** <https://www.linkedin.com/in/gohelvatsal/>

**GitHub Id: -** <https://github.com/vatsalgohel>

**Project Portfolio**

**Telecom Churn Case Study**

**Project Title:**

Enhancing Customer Retention Strategies through Predictive Analytics

**Problem Statement:**

Telecom companies face significant challenges in retaining customers due to high rates of churn, particularly influenced by service dissatisfaction, competitive pricing, and customer service issues. There is a need to develop effective predictive models to identify potential churners early and implement proactive retention strategies.

**Project Description:**

**Objective:**

Develop predictive models to forecast customer churn, prioritize at-risk customers, and recommend targeted retention strategies.

**Scope:**

Utilize historical customer data from a telecom company to train and validate machine learning models. Focus on logistic regression and random forest algorithms for their performance in churn prediction.

**Deliverables:**

Predictive models with high accuracy and recall for churn prediction, actionable insights for customer retention strategies.

**Methodology:**

**Tools and Technologies Used:**

Python, Scikit-learn, pandas, matplotlib, seaborn, Jupyter Notebook.

**Process:**

* Data preprocessing and exploratory data analysis (EDA).
* Feature engineering and selection.
* Model selection and hyperparameter tuning using cross-validation.
* Evaluation metrics: ROC-AUC, precision, recall, and confusion matrix.
* Implementation of logistic regression and random forest models.
* Comparative analysis of model performance on training and test datasets.

**Role and Responsibilities:**

* Data preprocessing, exploratory data analysis, and feature engineering.
* Model selection, hyperparameter tuning, and performance evaluation.
* Documentation of findings, insights, and recommendations.

**Results and Impact:**

**Outcome:**

Logistic regression and random forest models achieved ROC scores of 81.3% and 80.3% respectively on the test dataset, with recall rates of 76.03% and 74.51%.

**Impact:**

Improved ability to predict customer churn, enabling the telecom company to implement targeted retention strategies and reduce customer attrition rates.

**Challenges and Solution:**

**Challenges Faced:**

* Imbalanced dataset with a majority of non-churning customers.
* Optimal cutoff determination for balancing precision and recall.

**Solutions Implemented:**

* Applied oversampling techniques (SMOTE) to handle class imbalance.
* Utilized ROC curve analysis to determine optimal cutoff thresholds for maximizing recall.

**Conclusion and Learnings:**

**Summary:**

The project successfully developed and evaluated predictive models for customer churn prediction in the telecom industry. Logistic regression and random forest emerged as effective models for identifying potential churners.

**Learning Experience:**

Enhanced understanding of machine learning techniques for customer churn prediction, importance of feature engineering, model evaluation metrics, and the practical application of predictive analytics in business scenarios.

**Link for the file:**

<https://drive.google.com/drive/folders/1xA8V5Z8xk-307E7P1QoFJd89SynFCwjg?usp=sharing>