Customer Churn Prediction: A Comprehensive Guide

Table of Contents

1. Introduction:

In the dynamic world of business, understanding customer behavior is paramount. One of the key metrics businesses monitor is customer churn or attrition. This project delves into predicting potential customer churn using a variety of data points, enabling businesses to proactively address concerns and enhance customer retention.

2. Data Processing Steps:

2.1 Addressing Missing Data:

To maintain data integrity, missing values were either replaced using statistical imputation methods or omitted, ensuring a robust dataset for model training.

2.2 Categorical Data Transformation:

Fields like Gender and Location were transformed from categorical to numerical using LabelEncoder from scikit-learn, making them amenable for model training.

2.3 Normalizing Features:

Features with varying scales, such as Monthly Bill and Total Usage (GB), were standardized using StandardScaler, ensuring consistent data interpretation by the model.

3. Feature Engineering Techniques:

3.1 New Feature Creation:

Derived a feature 'Usage_to_Bill_Ratio' by dividing total usage by Monthly bill.

3.2 Selecting Relevant Features:

A combination of correlation analysis and feature importance metrics was employed to cherry-pick the most relevant features, ensuring the model's robustness and generalization capabilities.

4. Model Selection and Evaluation:

4.1 Choice of Model:

We opted for a Random Forest classifier, given its adeptness at handling intricate data patterns and its feature-ranking capabilities.

4.2 Performance Metrics:

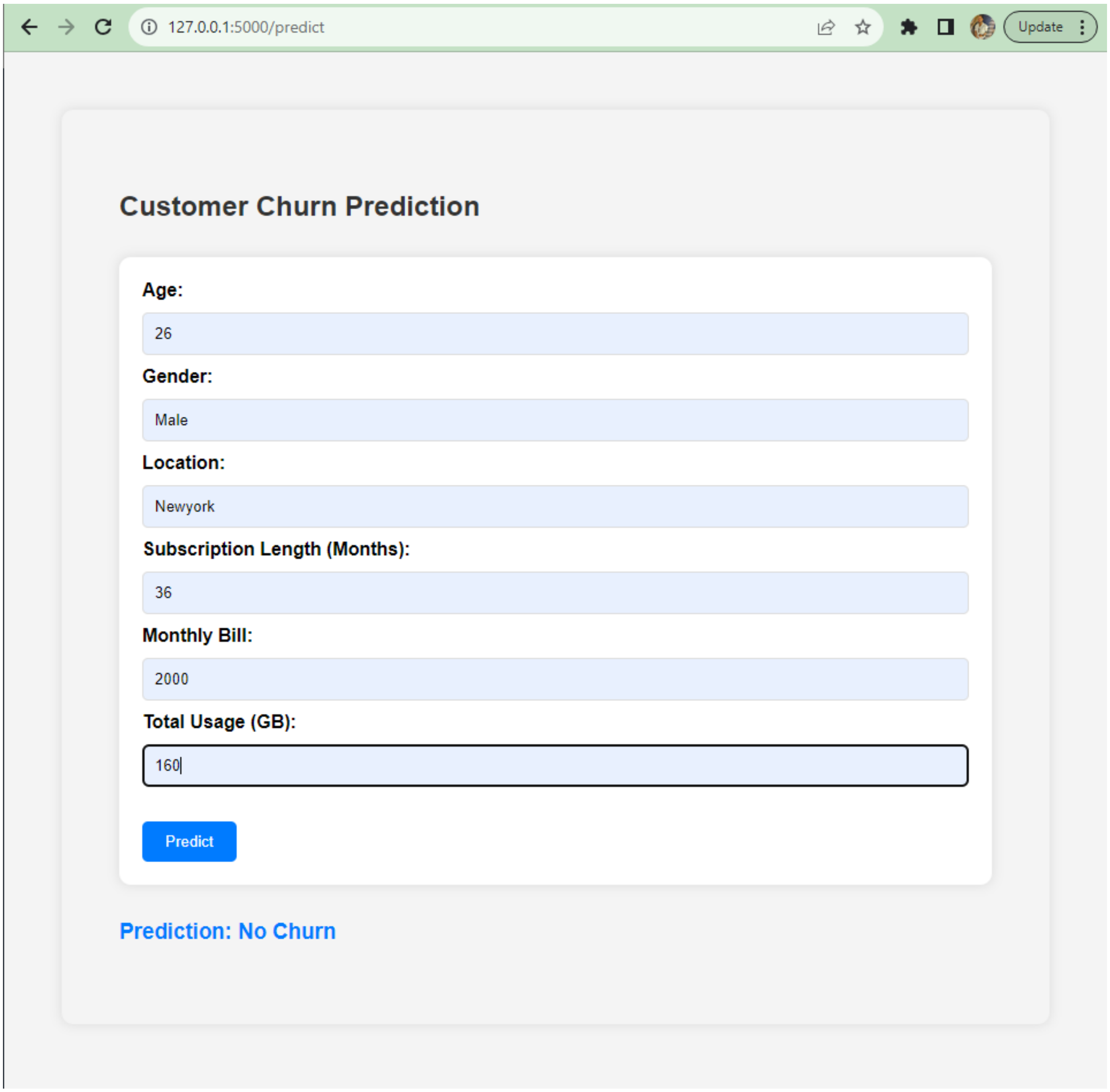
The efficacy of our model was assessed using a suite of metrics, including accuracy, F1-score, Precision and Recall. The model showcased an impressive accuracy of 49% on the validation dataset.

4.3 Refining Model Parameters:

To extract the best performance, hyperparameters were meticulously tuned using techniques like GridSearchCV.

5. Web Application Interface:

We've encapsulated our model within an intuitive web interface built using Flask, offering users the convenience of entering customer data and receiving instantaneous churn predictions.



6. Execution Instructions:

Secure a copy of the repository from (provide the link).

Transition to the project's directory.

Install the requisite libraries via pip install -r requirements.txt.

Initiate the Flask application with python app.py.

Access the web interface by navigating to http://127.0.0.1:5000 on your browser.

7. Software Dependencies:

This project leans on the following pivotal libraries:

Flask==2.0.1

numpy==1.21.0

pandas==1.3.0

scikit-learn==0.24.2

matplotlib==3.4.2

seaborn==0.11.1

gunicorn==20.1.0

8. Final Thoughts:

Our Customer Churn Prediction tool epitomizes the transformative power of machine learning in business decision-making. By pinpointing potential churn risks, businesses can recalibrate their strategies, fostering enhanced customer loyalty and driving sustained growth.