

# Personalized Healthcare Recommendations Project

## Objective

To develop a machine learning model that predicts donor behavior using a blood donation dataset and provides personalized healthcare recommendations to encourage continued donations or re-engage lapsed donors.

## Dataset Description

The dataset includes the following features:

- Recency: Months since last donation
- Frequency: Total number of donations
- Monetary: Total blood donated (removed due to redundancy)
- Time: Months since first donation
- Class: Target variable indicating donation status (1 = Donated, 0 = Not Donated)

## Data Preprocessing & Feature Engineering

Standardized numerical features using pipelines.

Dropped the 'Monetary' column due to its derivation from 'Frequency'.

Engineered a new feature:  $\text{Donation Rate} = \text{Frequency} / (\text{Time} + 1)$

## Model Selection and Training

Trained and compared multiple models:

- Logistic Regression
- Random Forest
- SVM
- XGBoost

Used pipelines for preprocessing + model training. Random Forest showed strong performance.

## Evaluation

# Personalized Healthcare Recommendations Project

Evaluated models using:

- Confusion Matrix
- Classification Report (Accuracy, Precision, Recall, F1-score)
- ROC Curve and AUC
- 5-fold Cross-validation for generalization performance

## Recommendation System

Implemented a function that maps model predictions to actionable recommendations:

- Class 1: Encourage continued donations
- Class 0: Suggest re-engagement strategies

## Deployment

Created a Flask API to accept donor data and return real-time personalized recommendations.  
Saved and reused the trained model using Pickle.

## Results

Random Forest achieved high accuracy and stable performance across cross-validation.  
The feature 'Donation Rate' contributed to better prediction of donor behavior.

## Future Work

- Apply hyperparameter tuning to improve model performance
- Use explainable AI (SHAP, LIME) to interpret predictions
- Add a simple user interface to the Flask app
- Incorporate additional demographic or health data for more comprehensive insights