D-Buddy: An R Shiny Web App for Diabetes Risk Prediction and Personalized Nutrition

- Rohit Dutta

Background & Motivation

Diabetes affects over 500 million people worldwide, with serious health risks if unmanaged. Early prediction and personalized diet are key to prevention. D-Buddy fills the gap by providing an intelligent, user-friendly platform for real-time risk assessment and tailored 3-course meal recommendations.

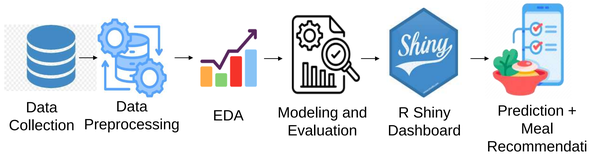
Objectives

- Predict diabetes risk using medical and lifestyle data
- Recommend personalized 3-course Veg/Non-Veg meals
- Apply analytics for health insights and decision-making
- Deploy an interactive R Shiny app for real-time use

Methods

We used the Pima Indians Diabetes dataset, applied preprocessing (missing value handling, normalization, encoding), and performed exploratory analysis to identify key predictors. After training multiple models, **XGBoost** was selected for its performance and deployed via a custom **R Shiny app** with integrated prediction and personalized meal planning.

Pipeline Diagram



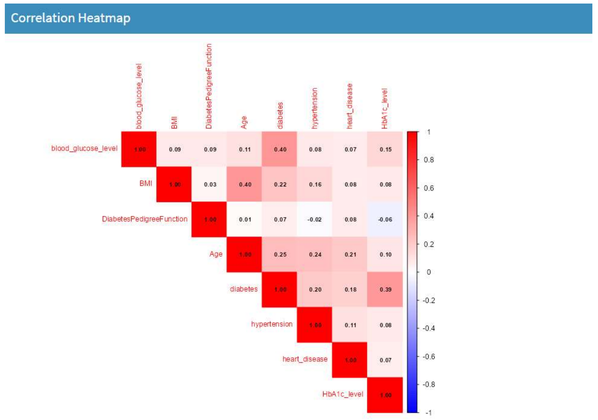
Meal recommendations are based on:

- **Glycemic Load (GL)** and **Glycemic Index (GI)** to ensure blood sugar-friendly meals
- **User's diabetes status** and **food preference** (Veg/Non-Veg)
- Balanced nutrition across breakfast, lunch, and dinner

Results

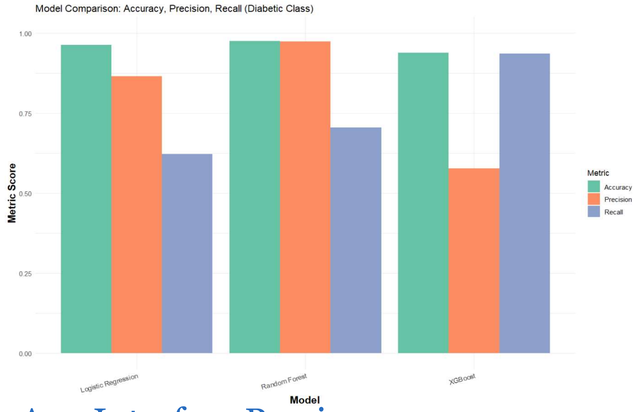
1. **Biivariate Plot: Distribution of Key Features:**
- **Glucose vs Diabetes**: Diabetic individuals mostly have glucose levels above **130 mg/dL**, often in the **140–200 mg/dL** range, while non-diabetics cluster below **120 mg/dL**.
- **HbA1c vs Diabetes**: Diabetics typically show **HbA1c > 6.5%**, commonly between **7–9%**, whereas non-diabetics are mostly below **6%**.

2. **Correlation Heatmap: Diabetes** shows moderate correlation with **Glucose (r = 0.40)**, **HbA1c (r = 0.39)**, and **Age (r = 0.25)**. Other features like **BMI** and **Hypertension** have weaker links (**r < 0.22**).**Glucose** and **HbA1c** emerge as the strongest predictors from the heatmap.

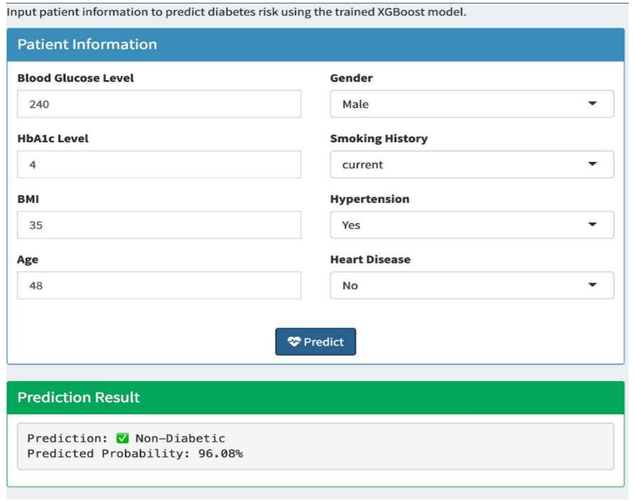


3. **Model Evaluation:** The XGBoost model demonstrated the highest **Recall (0.9361)** for the diabetic class, making it ideal for detecting positive cases, which is critical in medical diagnostics. While Random Forest achieved the best overall **Accuracy (0.9750)** and **Precision (0.9742)**, its lower Recall (0.7044) suggests it may miss more diabetic cases. XGBoost, with a balanced trade-off, was chosen for deployment to prioritize sensitivity in identifying diabetes risk.

Results



App Interface Preview



App Interface Preview

Personalized 3-Course Diet Plan

Meal	Food_Item	Category	GI	GL
Breakfast	Dhokla	Grains	35.00	7.00
Breakfast	Sprouts Salad	Legumes	25.00	3.75
Breakfast	Guava	Fruit	31.00	6.51
Lunch	Pesarattu	Grains	56.00	15.68
Lunch	Sweet Corn Soup	Vegetables	59.00	11.80
Lunch	Banana (Ripe)	Fruit	62.00	16.74
Dinner	Brown Rice	Grains	50.00	22.50
Dinner	Pav Bhaji	Vegetables	63.00	26.46
Dinner	Grapes	Fruit	59.00	8.85

Conclusions

D-Buddy successfully combines predictive analytics and personalized nutrition to support early diabetes detection and management. By leveraging XGBoost for accurate risk assessment and integrating intelligent meal recommendations, the R Shiny app offers a practical, user-friendly tool to promote healthier lifestyle choices and proactive diabetes prevention.

Explore the Project

Scan the QR code to access the full D-Buddy project on GitHub, including code, datasets, and the R Shiny app for diabetes prediction and meal planning.