Infosys®
Navigate your next

Predicting Obesity Levels Using Machine Learning and Deep Learning Methods

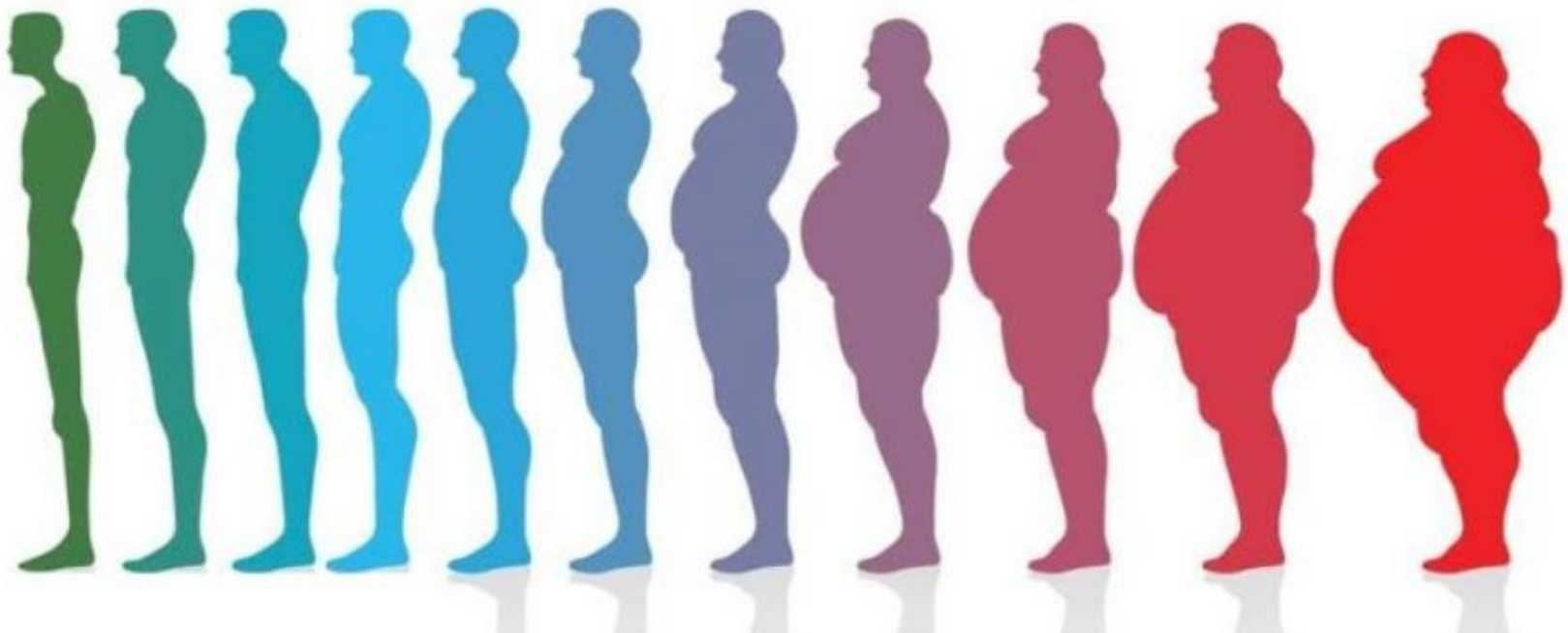
Infosys Springboard

Mentor: Mr. Narendra Kumar

Introduction

Problem Statement , Objective, Motivation & Problem Formulation

OBESITY IS NOW A GLOBAL EPIDEMIC!



Introduction
Problem Statement

- Obesity is a significant global health concern, with projections estimating that by 2030, 38% of the world's population will be overweight and 20% will be obese.
- High caloric intake, sedentary lifestyles, and transportation habits are identified as major contributors to this growing issue, according to the WHO.
- This condition is closely associated with an increased risk of chronic diseases, including diabetes, cardiovascular ailments, and certain types of cancer.
- Machine learning offers a promising alternative to traditional statistical methods for predicting obesity, enabling the development of personalized interventions and enhancing health policies.

Worldwide Increase in Obesity and Overweight

4.4 Million

Deaths by Obesity and Overweight

Increase of Obesity and Overweight in Adult and Children

40
Percentage of Overweight and Obese Adult in World

5
Is Number of Countries Able to Lower Down Obesity

15
Percentage of Adolescent Obesity

60
Percentage of Obese People in Developing Countries

Approximately 15% Of Obese People are from US, Globally

OBESITY AND OVERWEIGHT CONTRIBUTE TO:

- Cardiovascular Disease
- Diabetes
- Cancer
- Joint Pain

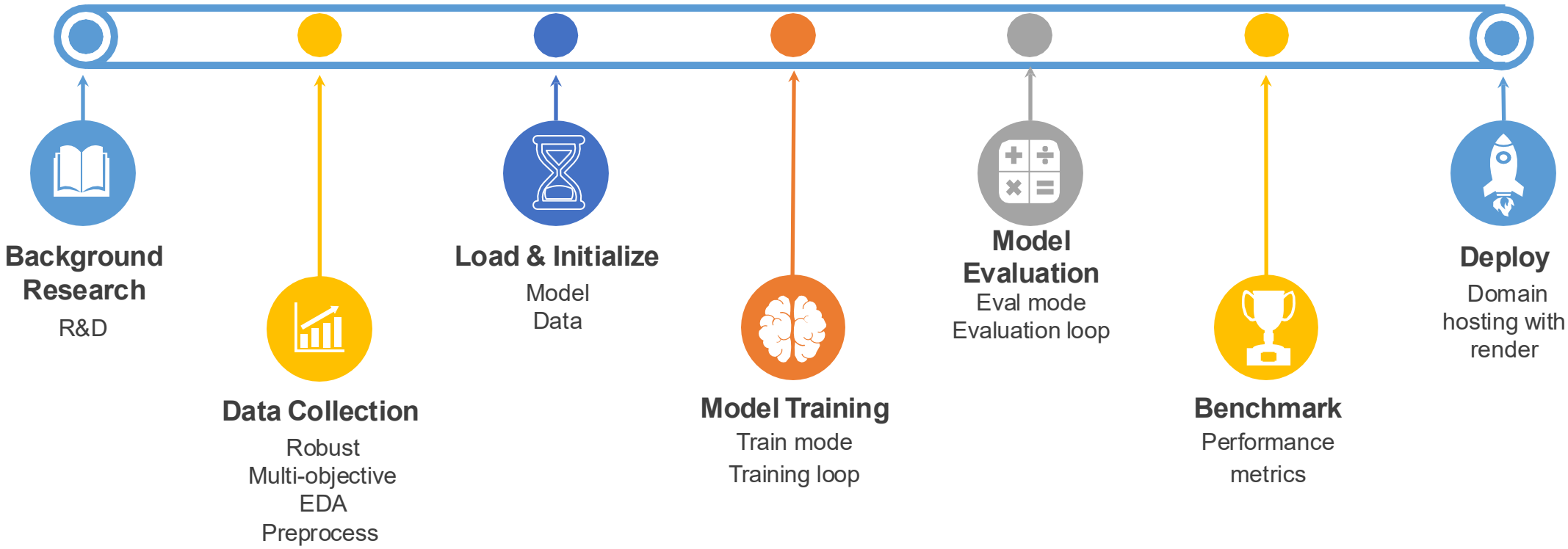
Motivation

- Promote early detection of obesity risks to enable timely lifestyle interventions.
- Deliver tailored insights for individuals to improve their health and well-being.
- Leverage machine learning to analyze lifestyle and physical condition data for accurate predictions.
- Make obesity prediction tools widely available to individuals and healthcare providers.
- Increase awareness about the impact of habits like diet and activity on obesity.

Objectives

1. Predict Obesity levels based on individuals' life styles
2. Help users identify risks of Obesity before hand
3. Provide an intuitive and user-friendly interface
4. Generate detailed report highlighting key factors
5. Educate users on how daily habits influence obesity levels

PROBLEM FORMULATION APPROACH



Proposed Methodology

Proposed Workflow

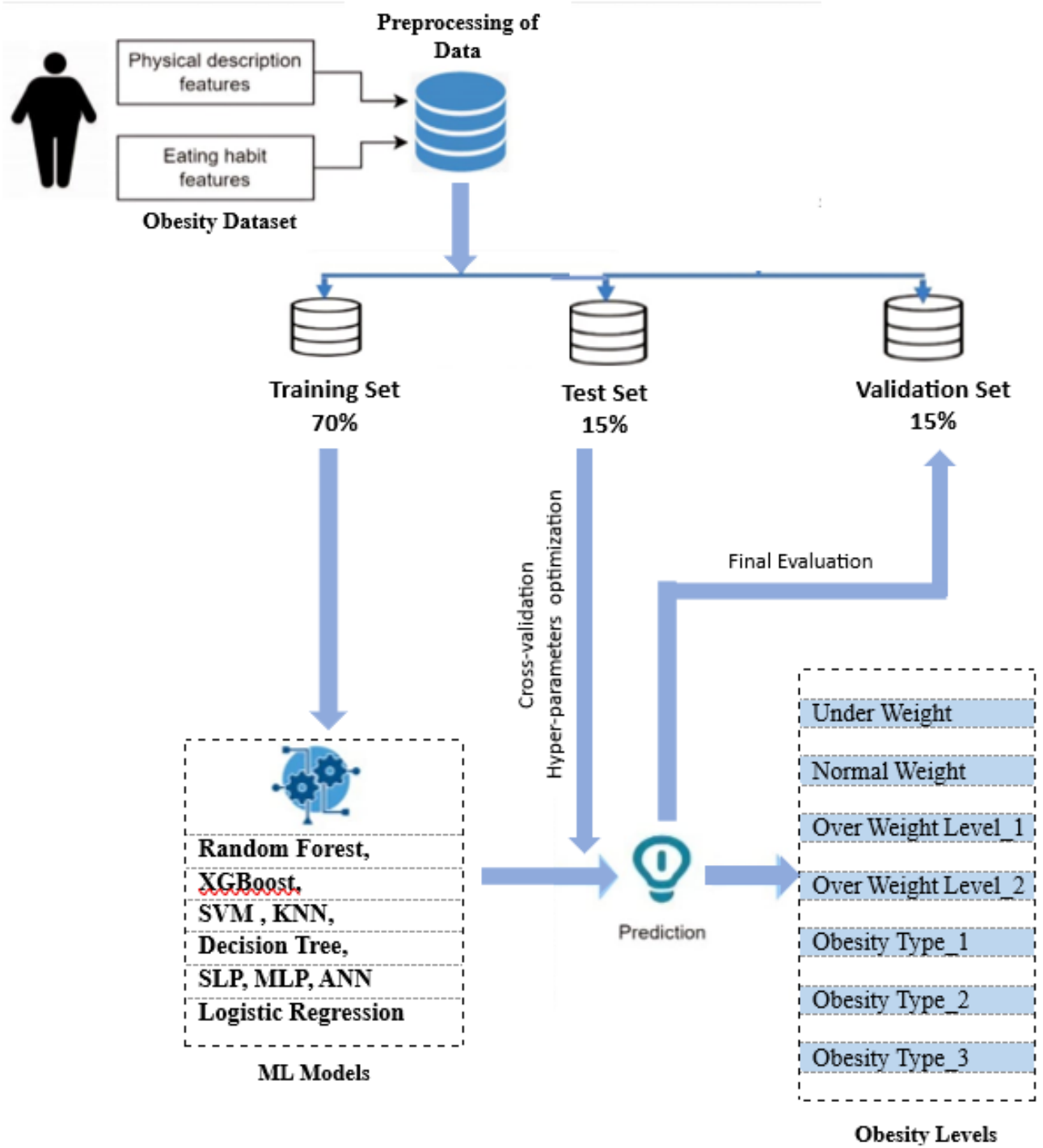


Fig.1 : Proposed Workflow for Obesity Level Prediction

Proposed methodology

Dataset

Source: Public clinical dataset containing obesity-related health and lifestyle data.

Size: Consists of 34,597 samples and 17 attributes.

Attributes: Covers demographics, health metrics, eating habits, lifestyle patterns, and obesity levels.

Data Characteristics: Includes numerical and categorical features with no missing values.

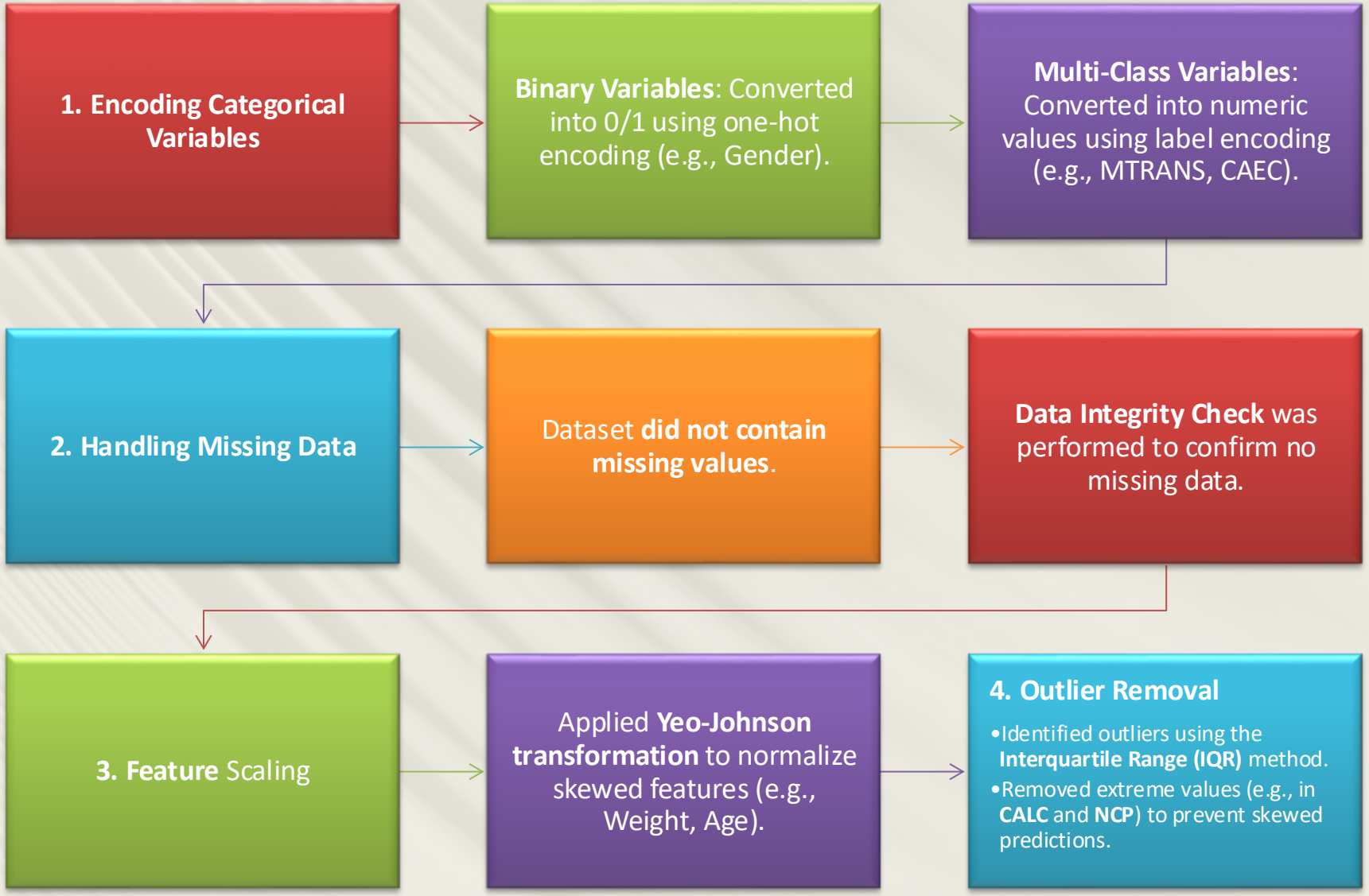
Insights Gained: Identified strong predictors like weight and physical activity influencing obesity classification.

Features

Gender: Feature, Categorical, "Gender"
Age : Feature, Continuous, "Age"
Height: Feature, Continuous
Weight: Feature Continuous
Family history with overweight: Feature, Binary, " Has a family member suffered or suffers from overweight? "
FAVC : Feature, Binary, " Do you eat high caloric food frequently? "
FCVC : Feature, Integer, " Do you usually eat vegetables in your meals? "
NCP : Feature, Continuous, " How many main meals do you have daily? "
CAEC : Feature, Categorical, " Do you eat any food between meals? "
SMOKE : Feature, Binary, " Do you smoke? "
CH2O: Feature, Continuous, " How much water do you drink daily? "
SCC: Feature, Binary, " Do you monitor the calories you eat daily? "
FAF: Feature, Continuous, " How often do you have physical activity? "
TUE : Feature, Integer, " How much time do you use technological devices such as cell phone, videogames, television, computer and others? "
CALC : Feature, Categorical, " How often do you drink alcohol? "
MTRANS : Feature, Categorical, " Which transportation do you usually use? "
NObeyesdad : Target, Categorical, "Obesity level"

Link to the Dataset

Data Preprocessing



Model Selection: 

Classical algorithms: Logistic Regression, Decision Tree, KNN, and SVM.

Ensemble methods: Random Forest and XGBoost.

Neural networks: ANN, MLP, and SLP.

Proposal
Model Training

Fig 2: Building , Training and Evaluation of the ML Models

Hyperparameter Tuning

Hyperparameters used :

- **XGBoost**: Tuning learning_rate, n_estimators, max_depth, subsample.
- **Random Forest**: Optimized n_estimators, max_depth, min_samples_split, min_samples_leaf.
- **Decision Tree**: Focused on max_depth, min_samples_split, min_samples_leaf.
- **Logistic Regression**: Tuned C, penalty, solver.
- **KNN**: Optimized n_neighbors, weights, metric.
- **SVM**: Adjusted C, kernel, gamma.
- **MLP**: Tuned hidden_layer_sizes, activation, learning_rate_init.
- **ANN**: Similar to MLP but with additional layers for more complex data relationships.

Evaluation Metrics

Accuracy Score: Percentage of correctly predicted obesity levels.

Classification Report: Provided metrics such as Precision, Recall, and F1-score for each obesity class.

Confusion Matrix: Visualized the number of correct and incorrect predictions across all classes.

Comparative Analysis

Sr no.	Approach	ALGORITHM USED	Accuracy (Before Tuning)	Accuracy (After Tuning)
1.	XGBOOST	GridSearchCV	97.35	97.49
2.	RANDOM FOREST	Bayesian Optimization	96.52	96.93
3.	DECISION TREE	GridSearchCV	95.82	96.52
4.	SLP	GridSearchCV	95.82	96.11
5.	SVM	GridSearchCV	95.96	96.10
6.	LOGISTIC REGRESSION	RandomizedSearchCV	95.13	95.54
7.	MLP	RandomizedSearchCV	94.29	94.57
8.	ANN	RandomizedSearchCV	93.18	94.29
9.	KNN	GridSearchCV	89.01	91.37

17

Results

Results

XGBOOST METRICS

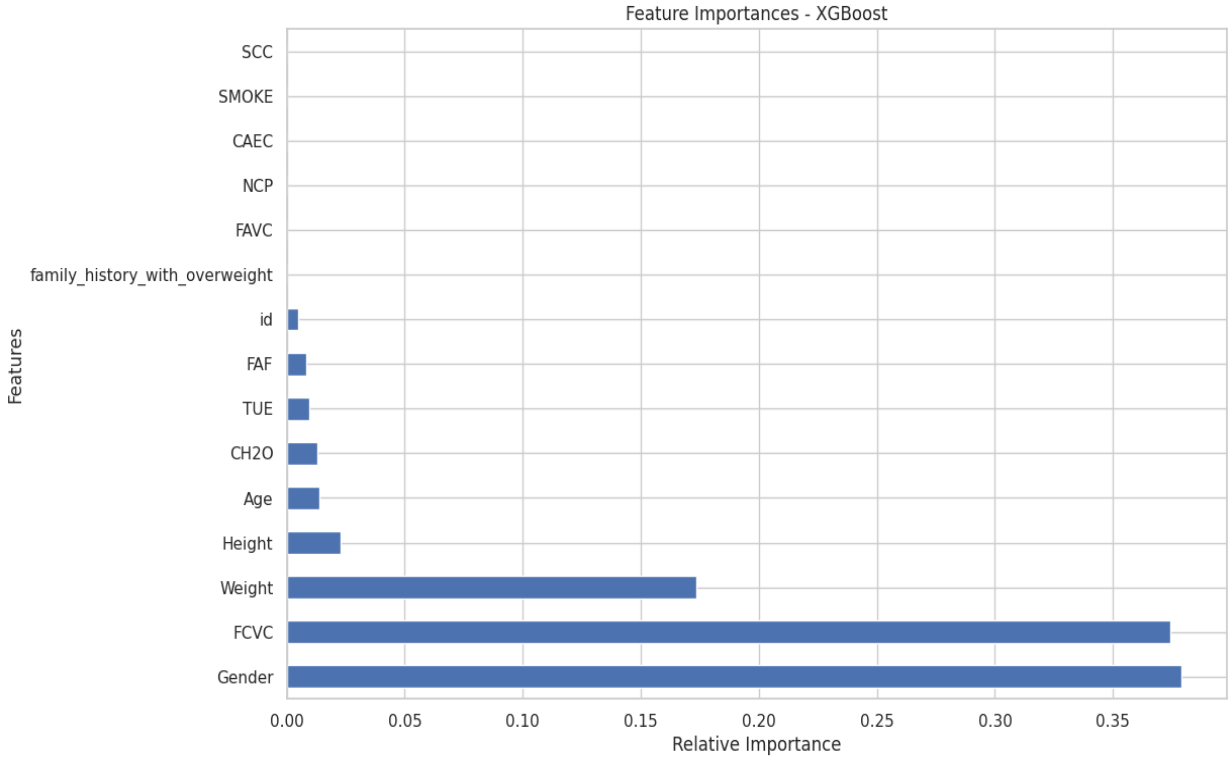


Fig.3: Feature importance analysis for XGBoost Model.

Classification Report:

	precision	recall	f1-score	support
0	0.93	1.00	0.96	13
1	0.91	0.95	0.93	42
2	0.96	0.93	0.95	56
3	0.98	0.98	0.98	116
4	1.00	1.00	1.00	426
5	0.83	0.85	0.84	40
6	0.92	0.85	0.88	26
accuracy			0.97	719
macro avg	0.93	0.94	0.93	719
weighted avg	0.98	0.97	0.97	719

Accuracy: 0.9749652294853964

Fig.4: Classification report of the XGBoost model.

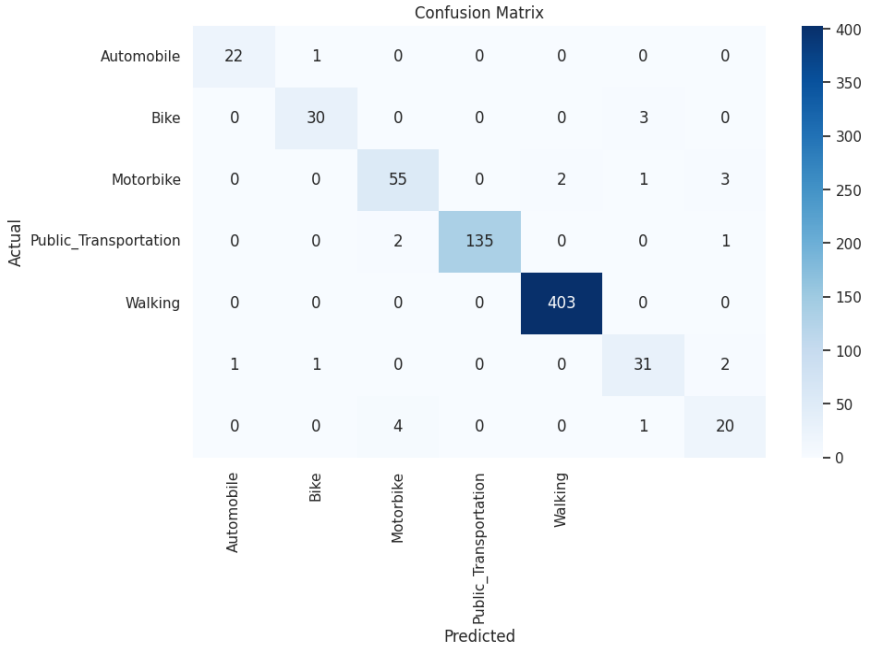
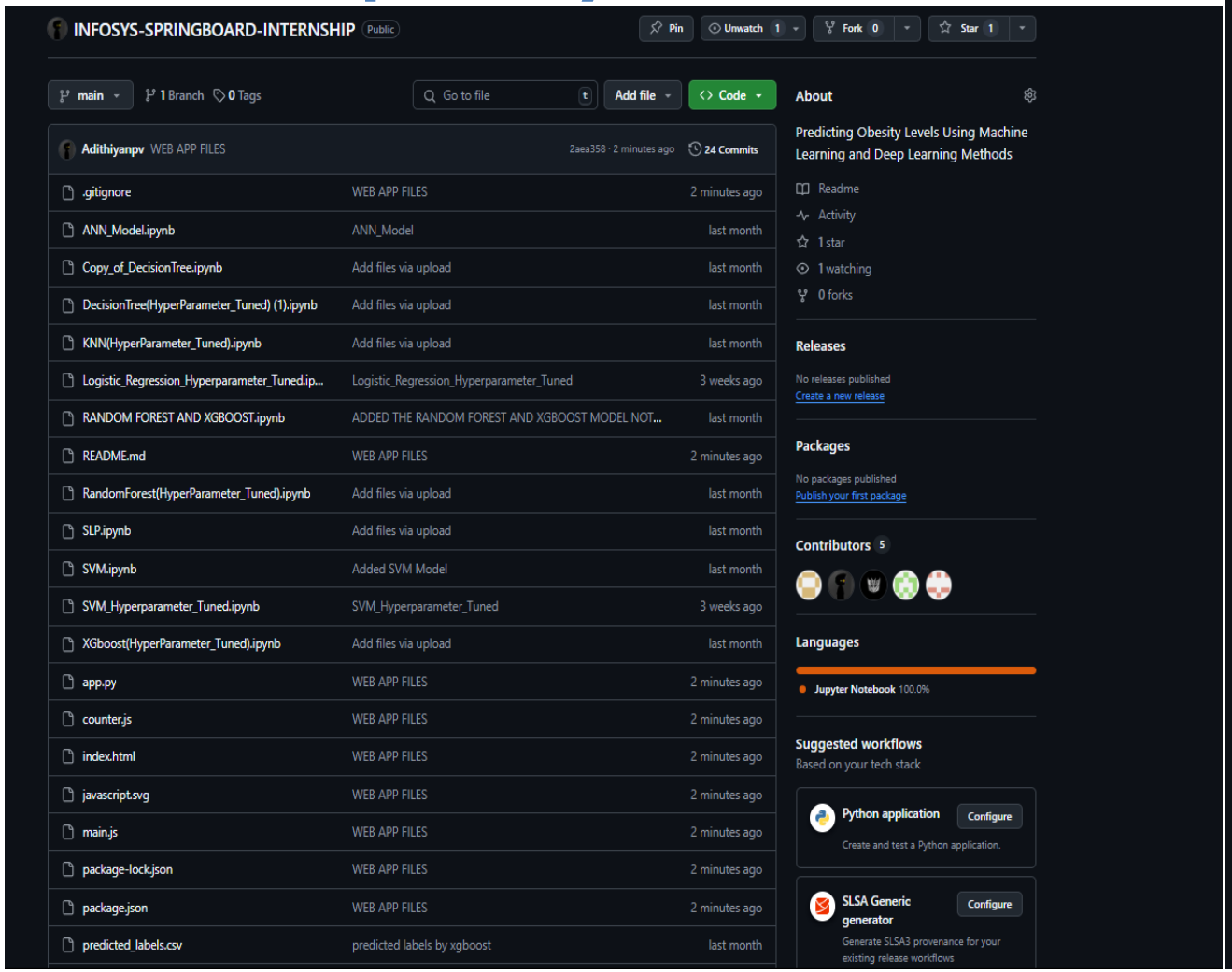


Fig.5: Confusion matrix of XGBoost model.

Results

GitHub Repository



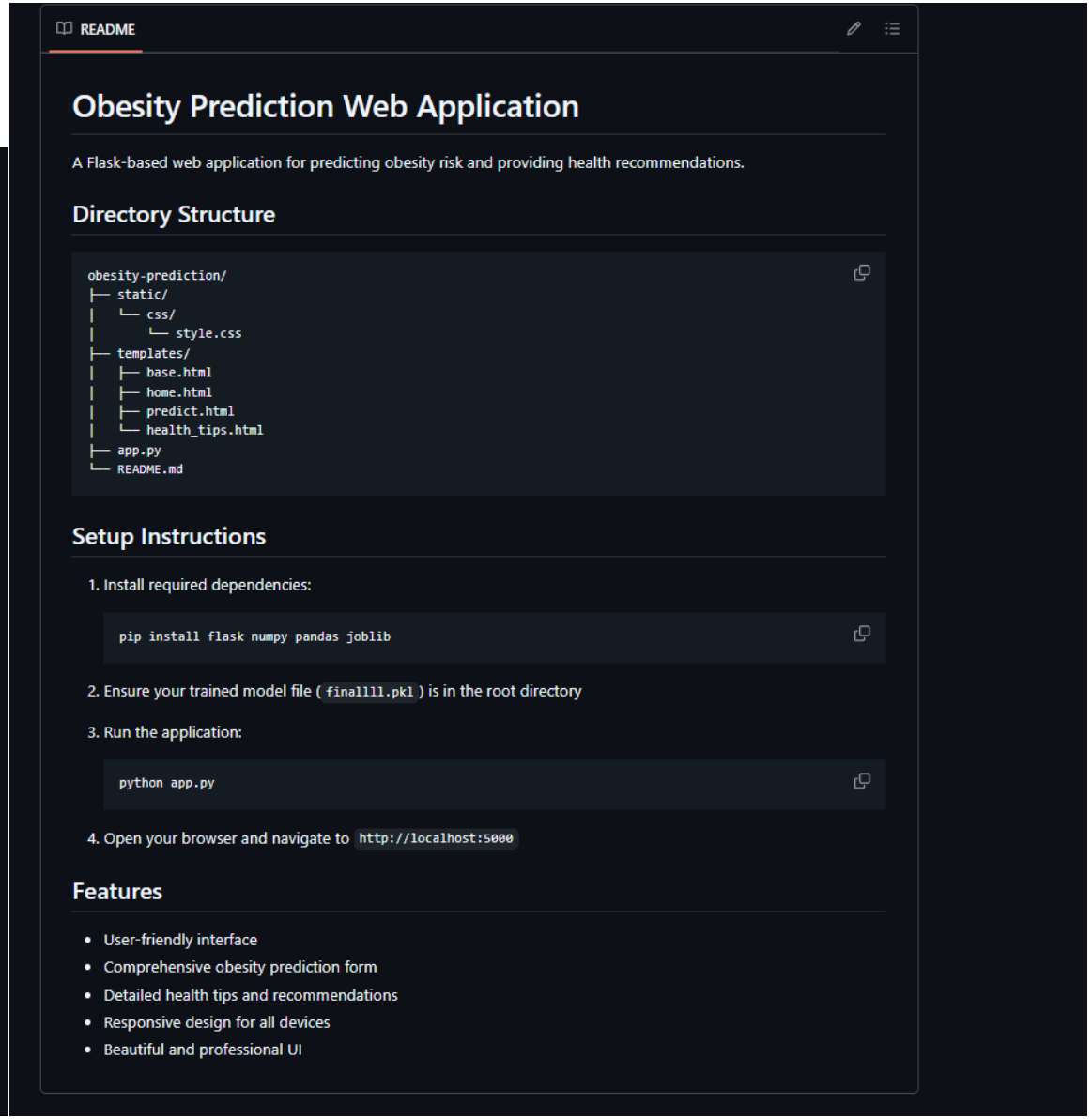


Fig.6 : Screenshot of the GitHub repository.

Infosys_Springboard_Obesity_Level_Prediction

Interface of our Web Application

What is Obesity?

Obesity is a medical condition where an individual has excessive body fat, which can lead to increased risks for many diseases, including diabetes, heart disease, and more.

Understanding your obesity level is important for taking preventive actions and managing your health effectively.

Check Your Obesity Level

Select Your Prediction Model

Choose a model to predict obesity based on different categories. Each model provides varying degrees of prediction granularity. Select the one that best suits your needs.

4-Category Classification

Classifies individuals into 4 categories: Underweight, Normal weight, Overweight, and Obesity.

Select this Model

7-Category CLassification

Classifies individuals into 7 categories: Insufficient Weight, Normal Weight, Obesity Type_I, Obesity Type II, Obesity Type III, Overweight Level I, and Overweight Level II.

Select this Model

Back to Information Page

Obesity Risk Level Assessment Web Application

(Obesity Level - 7 Category Classification)

Gender:
Male

Age:
20

Height (in m):
1.75

Weight (in kg):
98

Family History with Overweight:
No

Frequent Consumption of High Caloric Food:
Yes

Physical Activity Frequency (days per week):
2

Time Using Technology Devices (hours):
12

Frequency of Vegetables Consumption (1-3):
3

Number of Main Meals in a Day:
2

Consumption of Food Between Meals:
Sometimes

Smoking Habit:
No

Daily Water Consumption (L):
3

Calories Monitoring Daily:
Yes

Alcohol Consumption:
No

Mode of Transportation:
Bike

Get Obesity Level | Reset

Go Back to Home Page

Obesity Level Result

Your 4 Category Obesity Level Result:

Normal

Suggestions for You!! 😊

Diet Plan

Here is a personalized diet plan for you based on your BMI category 'Normal':

- Aim to consume a balanced diet with a variety of fruits, vegetables, whole grains, lean proteins, and healthy fats.
- Focus on portion control to maintain your weight within the normal range.
- Stay hydrated by drinking plenty of water throughout the day.
- Limit processed foods, sugary beverages, and high-fat snacks.
- Incorporate regular physical activity into your routine to support overall health and well-being.
- Consider consulting with a nutritionist or dietitian for personalized guidance on meal planning and dietary choices.

Exercise Plan

Here is a personalized exercise plan based on your BMI category 'Normal':

- Aim to engage in at least 150 minutes of moderate-intensity aerobic exercise per week, such as brisk walking, cycling, or swimming.
- Incorporate strength training exercises at least 2 days a week to build and maintain muscle mass.
- Include activities that you enjoy, such as dancing, yoga, or playing a sport, to make exercise more enjoyable and sustainable.
- Set specific and achievable fitness goals to stay motivated and track your progress over time.
- Remember to listen to your body, stay hydrated, and consult with a fitness professional if you have any specific health concerns or conditions.

Go Back to Home Page | Check Obesity Level Again

Results Recording of our Web Application

What is Obesity?

Obesity is a medical condition where an individual has excessive body fat, which can lead to increased risks for many diseases, including diabetes, heart disease, and more.

Understanding your obesity level is important for taking preventive actions and managing your health effectively.

Check Your Obesity Level

Results

Deployment

- **Visit our deployed web-app at:**

https://predicting-obesity-levels.onrender.com

THANK YOU !!!