**Problem Statement :** To predict obesity levels using health and lifestyle attributes like Age, Weight, Height, and eating/activity habits through a Random Forest model, deployed via a Streamlit app for real-time predictions.

**Objective**

To develop a machine learning model that can accurately classify individuals into various obesity levels based on their demographic characteristics and lifestyle habits. The model aims to support healthcare practitioners in early identification of obesity trends and facilitate proactive health management.

**Dataset Features**

The model uses a dataset (Obesity prediction.csv) that typically includes the following types of features:

* Age
* Gender
* Height, Weight
* Eating habits (frequency of fast food, vegetable consumption, etc.)
* Physical activity (daily activity level, time spent exercising, etc.)
* Lifestyle habits (smoking, alcohol consumption)
* Family history and other health indicators

**Target Variable:**

* Obesity (Categorical values such as Underweight, Normal Weight, Overweight, Obese, etc.)

**Approach**

**1. Data Preprocessing**

* Handle categorical variables using Label Encoding.
* Standardize numerical data using StandardScaler.

**2. Splitting the Data**

* Train-test split (80/20) for evaluation.

**3. Model Training**

* Trained using RandomForestClassifier with class\_weight='balanced' to handle imbalanced classes.

**4. Prediction**

* Input data is preprocessed and passed to the trained model to get an encoded class prediction.
* Prediction is decoded back to human-readable obesity labels.

**5. Evaluation**

* Accuracy
* Precision (Weighted)
* Recall (Weighted)

**Model Output**

The model returns the **predicted obesity level** and supports performance evaluation using common classification metrics.

**Key Highlights**

* Fully automated preprocessing and encoding
* Easy integration and extensibility
* Well-suited for real-world health monitoring systems
* Modular code with reusable components

**Applications**

* **Healthcare & Clinics**: Use as a screening tool during routine health checkups.
* **Wellness Programs**: Targeted recommendations based on predicted obesity category.
* **Research**: Analyze obesity patterns across populations.

**Sample Usage Code**

python

CopyEdit

from model import ObesityModel

# Load and train the model

model = ObesityModel()

# Predict using new input

sample\_input = [value1, value2, ..., valueN] # Replace with actual feature values

prediction\_encoded = model.predict(sample\_input)

# Decode prediction to label

label = model.get\_target\_encoder().inverse\_transform([prediction\_encoded])[0]

print("Predicted Obesity Class:", label)

# View evaluation metrics

metrics = model.get\_metrics()

print("Model Performance:", metrics)

**Folder Structure**

bash

CopyEdit

📁 Obesity-Prediction/

├── model.py # Core ML logic

├── Obesity prediction.csv # Input dataset (add manually)

├── README.md # Usage and setup

├── Project\_Documentation.md # Full project description

**Contact**

For questions or feedback, feel free to open an issue or pull request.

**License**

This project is intended for educational and non-commercial use only.

**Outputs :**



