OBESITY Category Prediction

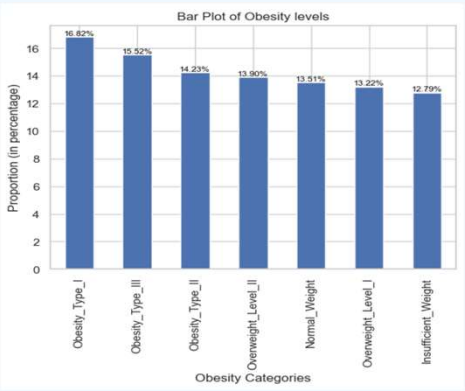
Introduction

Obesity is a growing concern worldwide, and its prediction is crucial for preventing and managing related health problems. obesity is one of the biggest risk factors for a variety of chronic diseases, including heart disease and cancer. The World Health Organization (WHO) defines obesity as an abnormal or excessive deposition of fat that has the potential to severely impact health. Obesity can have a detrimental impact on health (BMI).

Obesity categories

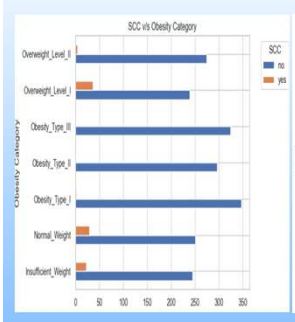
Obesity categories classify individuals based on their body mass index (BMI) into different levels of weight status, ranging from underweight to severe obesity. These categories help to assess health risks associated with weight and tailor appropriate interventions for prevention and treatment.

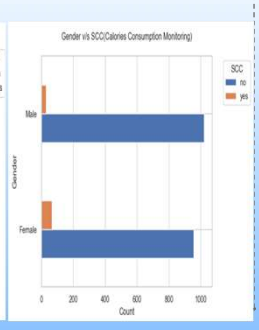
- **Obesity_Type_I**: This category has the highest proportion among all categories, indicating severe obesity.
- **Obesity_Type_III**: Another severe obesity category with a high proportion, though slightly lower than Obesity_Type_I.
- **Obesity_Type_II**: This category represents moderate obesity, with a lower proportion compared to Types I and III.
- **Overweight_Level_II**: Individuals in this category are classified as significantly overweight but not yet in the obesity range.
- **Normal_Weight**: This category indicates a healthy weight range and falling within the normal range.
- **Overweight_Level_I**: Individuals in this category are slightly overweight, closer to the normal weight range.
- **Insufficient_Weight**: This category represents individuals with insufficient weight, potentially underweight.



Comparison of Gender with SCC variable (Monitoring Consumption of Calories)

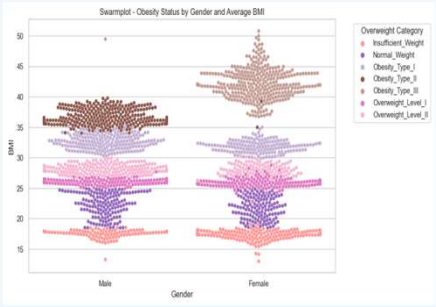
Females are more likely to monitor calorie consumption. Individuals who do not monitor their calorie intake are more likely to have either Obesity Type I or Obesity Type III. There is a statistical significant correlation between Gender and SCC at 0.05 significance level.

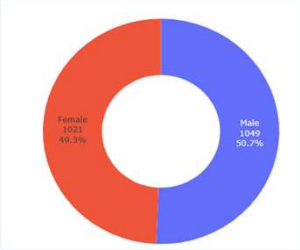




Comparison of Gender with Average BMI against Obesity levels

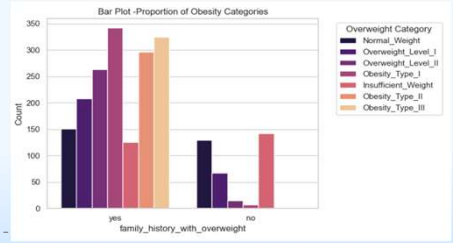
- Male tend to have higher rate of Obesity Type II as compared to Females who are more prone to Obesity Type III .
- 51% respondents are male and 49% are female. Out of Male proportion 76% respondents are overweight and obese and maximum have Obesity Type II whereas in female 70% of the total female respondents are overweight and obese. Maximum females belongs to Obesity Type 3.

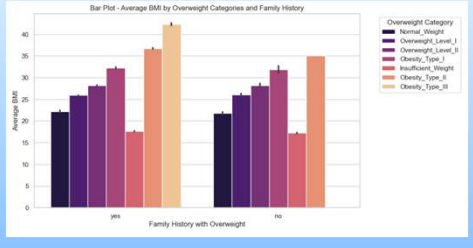




Comparison of Family History with Overweight and Obesity Categories by Average BMI

Individuals with no family history tend to have lower rates of obesity across all categories compared to those with a family history. However, those with a family history, particularly of obesity, exhibit higher rates of obesity, with Obesity Type III having the highest prevalence. This suggests a potential genetic influence on obesity risk.

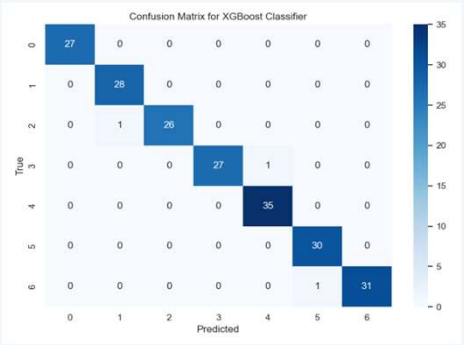




Results

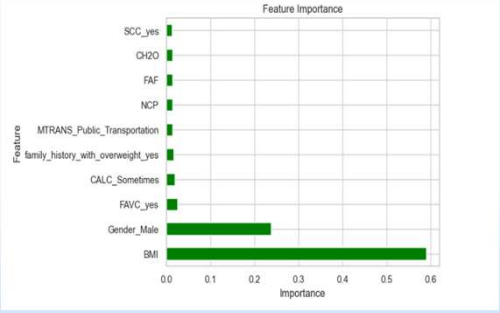
XGBClassifier demonstrates the best overall performance among and here's the reasoning:
1.Highest Accuracy: XGBClassifier achieves the highest accuracy of 99%, indicating superior overall predictive performance.
2.Balanced F1-score: It maintains a high weighted average F1-score of 0.99, reflecting its ability to balance precision and recall effectively.
3.Robust Recall: XGBClassifier exhibits strong weighted average recall of 0.99, ensuring it captures a high proportion of positive instances across all classes.
4.Precise Predictions: With a weighted average precision of 0.99, XGBClassifier demonstrates precision in its positive predictions across all classes.
5.Overall Superiority: Considering accuracy, F1-score, recall, and precision, XGBClassifier emerges as the top-performing model, offering reliable and balanced predictions for the dataset.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	27
1	0.97	1.00	0.98	28
2	1.00	0.96	0.98	27
3	1.00	0.96	0.98	28
4	0.97	1.00	0.99	35
5	0.97	1.00	0.98	30
6	1.00	0.97	0.98	32
accuracy			0.99	207
macro avg	0.99	0.99	0.99	207
weighted avg	0.99	0.99	0.99	207



Feature Importance

The top 10 features demonstrate varying degrees of importance in predicting the target variable. BMI emerges as the most influential factor, followed by gender, high-caloric food consumption, and alcohol consumption. Family history of overweight and transportation mode hold lesser significance, while variables like number of main meals, physical activity frequency, water consumption, and calorie monitoring contribute minimally to the model's predictive power.



Conclusion

XGBoost model for predicting obesity levels demonstrates promising performance, with key features such as BMI and gender playing significant roles in classification. Additionally, lifestyle factors like high-caloric food consumption and family history of overweight contribute to the predictive accuracy. While the model provides valuable insights, further refinement and validation may enhance its effectiveness in real-world applications.