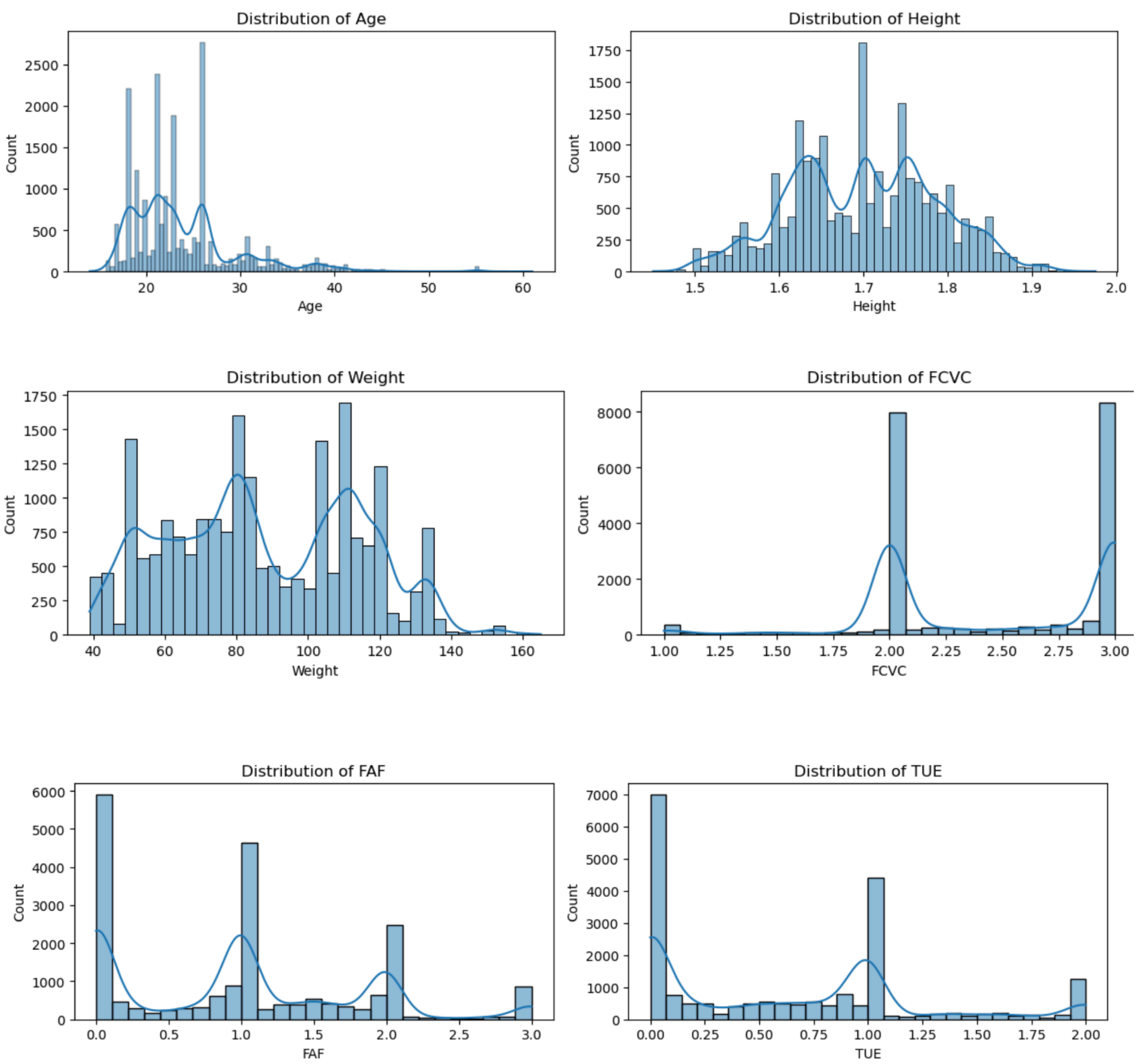
OBESITY CLASSIFICATION

Data Loading and Exploration

- The dataset contains **20,758 entries** and **23 columns** after preprocessing, including both categorical and numerical features. Some columns were transformed, and unnecessary ones were dropped to optimize the dataset for further analysis.
- **Missing Values**: There are no missing values in the dataset, ensuring that no further imputation or handling of missing data is required.
- Column Types:
    - The dataset contains **binary features** such as family history of obesity, FAVC (frequent consumption of high-calorie food), and SMOKE, which were converted to boolean types for clarity.
    - **Categorical variables** (like eating habits, alcohol consumption, and transportation methods) were one-hot encoded for compatibility with machine learning algorithms.
- Statistical Summary:
    - The dataset provides a range of numerical features like **Age**, **Height**, **Weight**, and various dietary and physical activity factors.
    - The average age is around **23.84 years**, with a standard deviation of **5.69 years**, showing that the data spans multiple age groups.
    - The **average height** is **1.7 meters**, and the **average weight** is **87.89 kg**, indicating a diverse range of body measurements within the dataset.
    - FCVC (Frequency of Vegetable Consumption) and NCP (Number of Meals per Day) show most values around 2–3, while CH2O (Water

Consumption) averages around 2 liters per day, highlighting lifestyle habits.

Data Analysis and Visualization of Obesity Levels



The insights from the distribution plots of numerical variables are:

1. **Age Distribution**:

   The majority of individuals are concentrated between the ages of 15 and 30.

There's a significant skew towards younger individuals, with very few data points beyond the age of 40, indicating a predominantly younger population in the dataset.

2. **Height Distribution**:

   The height distribution shows a peak around the 1.6-1.7 meter range, with a normal-like distribution. This suggests that the sample contains a typical variation in height, with most individuals in the average height range.

3. **Weight Distribution**:

   The weight distribution is multimodal, with distinct peaks around 60 kg, 80 kg, and 100 kg. This indicates that the dataset captures a broad range of individuals across different weight categories, useful for analyzing obesity levels.

4. **FCVC (Frequency of Consumption of Vegetables)**:

   This variable shows a high concentration at two specific points: 2 and 3. Most individuals report consuming vegetables frequently, which is a crucial dietary factor in assessing health and obesity risks.

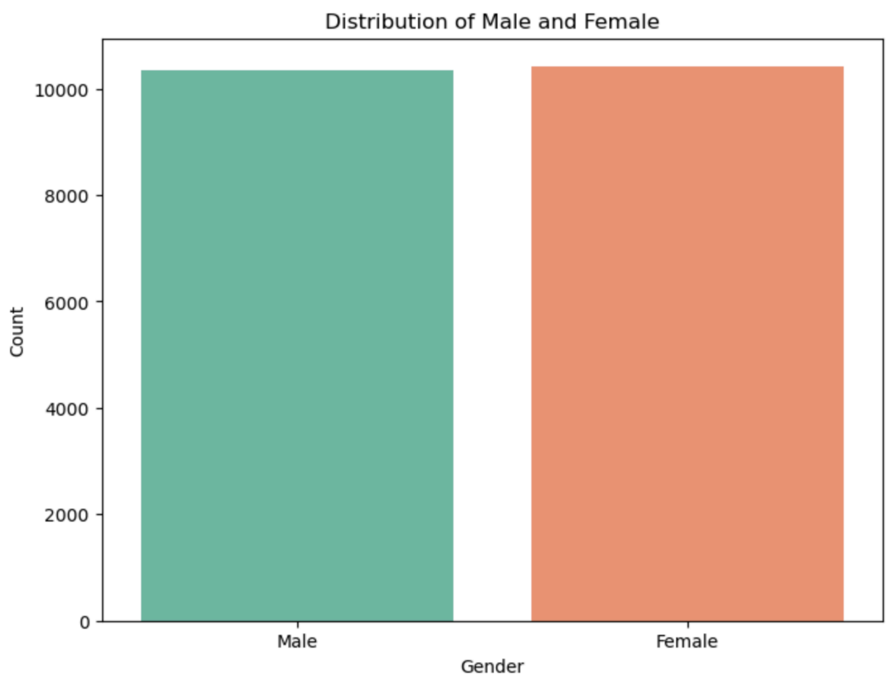
5. **FAF (Physical Activity Frequency)**:

   The majority of individuals report low physical activity, with a large number of data points near zero. This could be a significant indicator of sedentary lifestyle trends within the population, which correlates with obesity risks.

6. **TUE (Time using technology for entertainment)**:

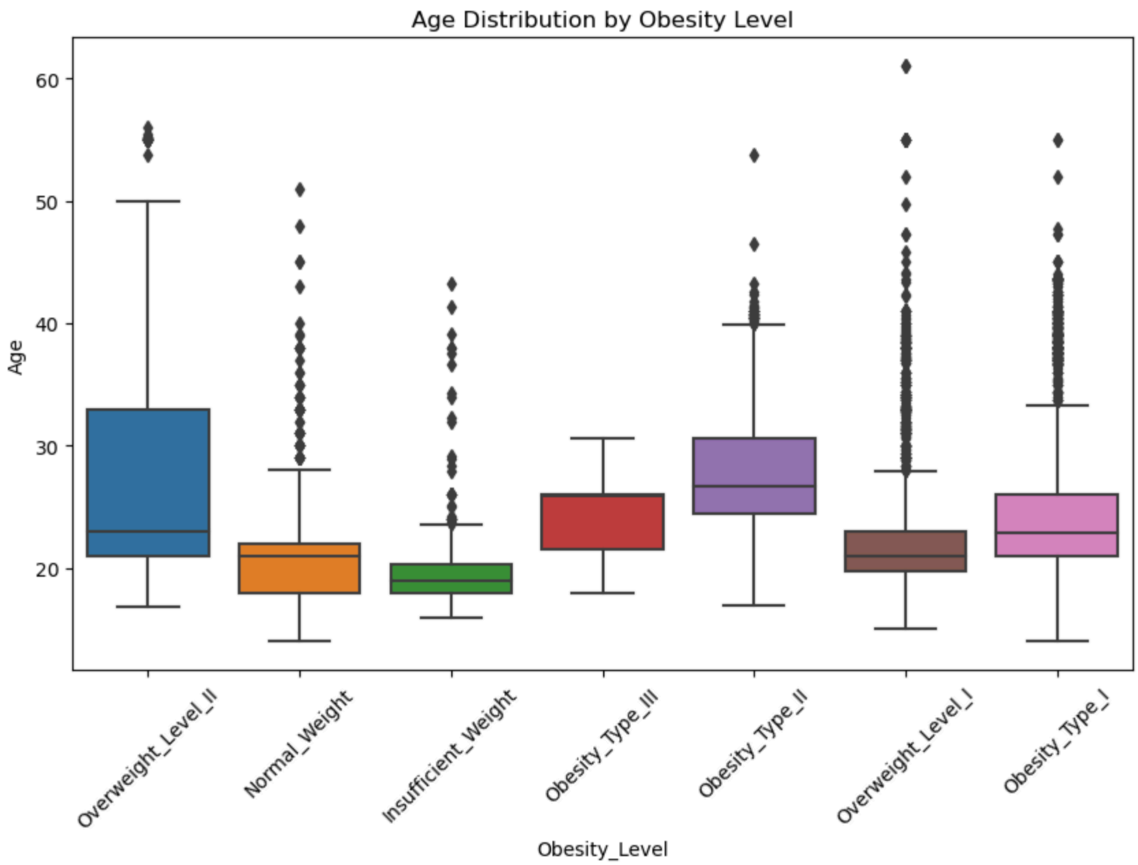
   There is a significant skew towards lower usage, but a small peak around 1-2 hours. This suggests that most individuals spend relatively little time on technology for entertainment, though a notable portion engages with it for a moderate amount of time.

Gender Distribution in the Dataset:



The gender distribution in this dataset is well-balanced, with nearly equal representation of both male and female participants. This balance ensures that the machine learning models developed in this project are less likely to exhibit gender bias, providing a more accurate comparison of obesity risk factors across genders. The equal distribution of genders also allows for better generalization of the models, reducing the likelihood of skewed predictions due to gender imbalance.

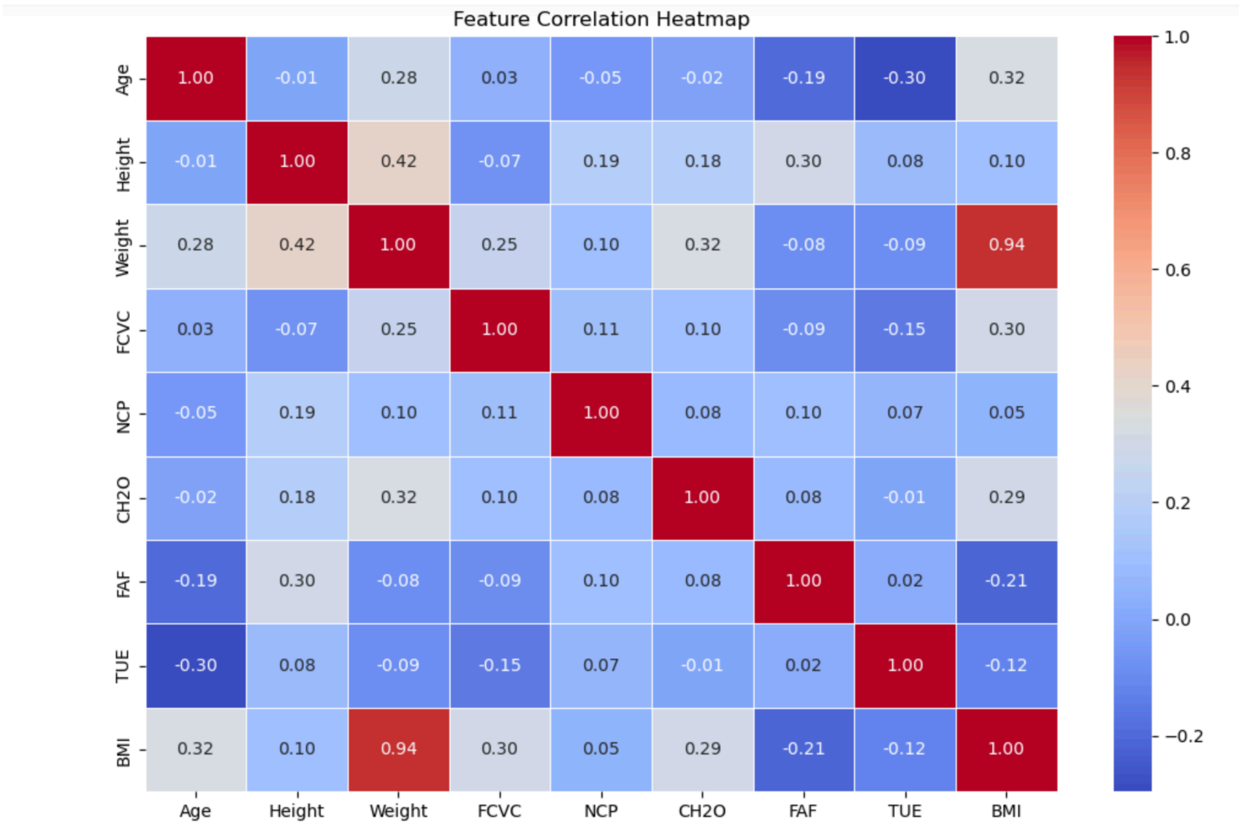
Age Distribution Across Obesity Levels:



- Individuals classified as Obesity_Type_III and Obesity_Type_II tend to be older, with a wider age range compared to those in the Normal_Weight or Insufficient_Weight categories.
- The Overweight_Level_II group shows a broader spread in age, suggesting that age may not correlate strongly with this specific level of obesity.
- Individuals in the Normal_Weight and Insufficient_Weight categories are generally younger, with a more concentrated age range.

These insights suggest that age plays a role in obesity classification, with higher obesity levels being more frequent among older individuals, whereas younger individuals are more likely to fall within Normal_Weight or Insufficient_Weight categories.

Feature Correlation Heatmap:



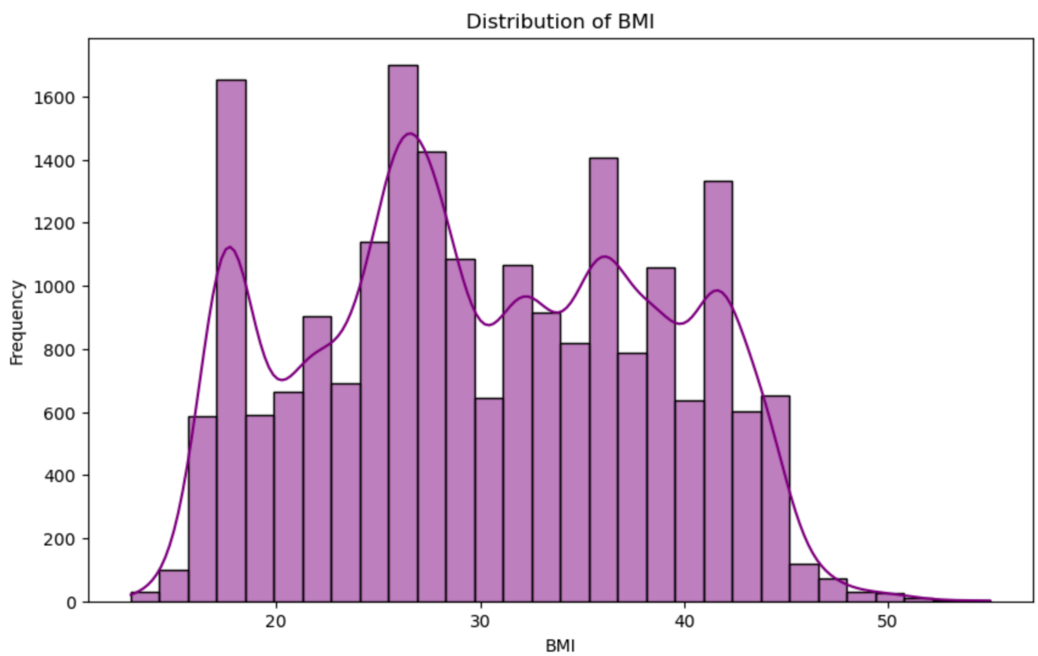
The correlation heatmap provides a clear view of the relationships between numerical features in the dataset. Some key insights from this heatmap include:

- **BMI and Weight** show a strong positive correlation (0.94), which is expected since BMI is calculated based on weight and height.
- **Age** has a moderate positive correlation with BMI (0.32) and weight (0.28), suggesting that older individuals tend to have higher BMI and weight.
- **FAF (Frequency of Physical Activity)** is weakly negatively correlated with BMI (-0.21), indicating that more frequent physical activity may contribute to a lower BMI.

- **TUE (Time Using Technology)** has a weak negative correlation with BMI (-0.12), suggesting a minor relationship between increased screen time and a lower BMI.
- **Height and Weight** have a moderate positive correlation (0.42), showing that taller individuals tend to weigh more, though it doesn't directly indicate obesity levels.

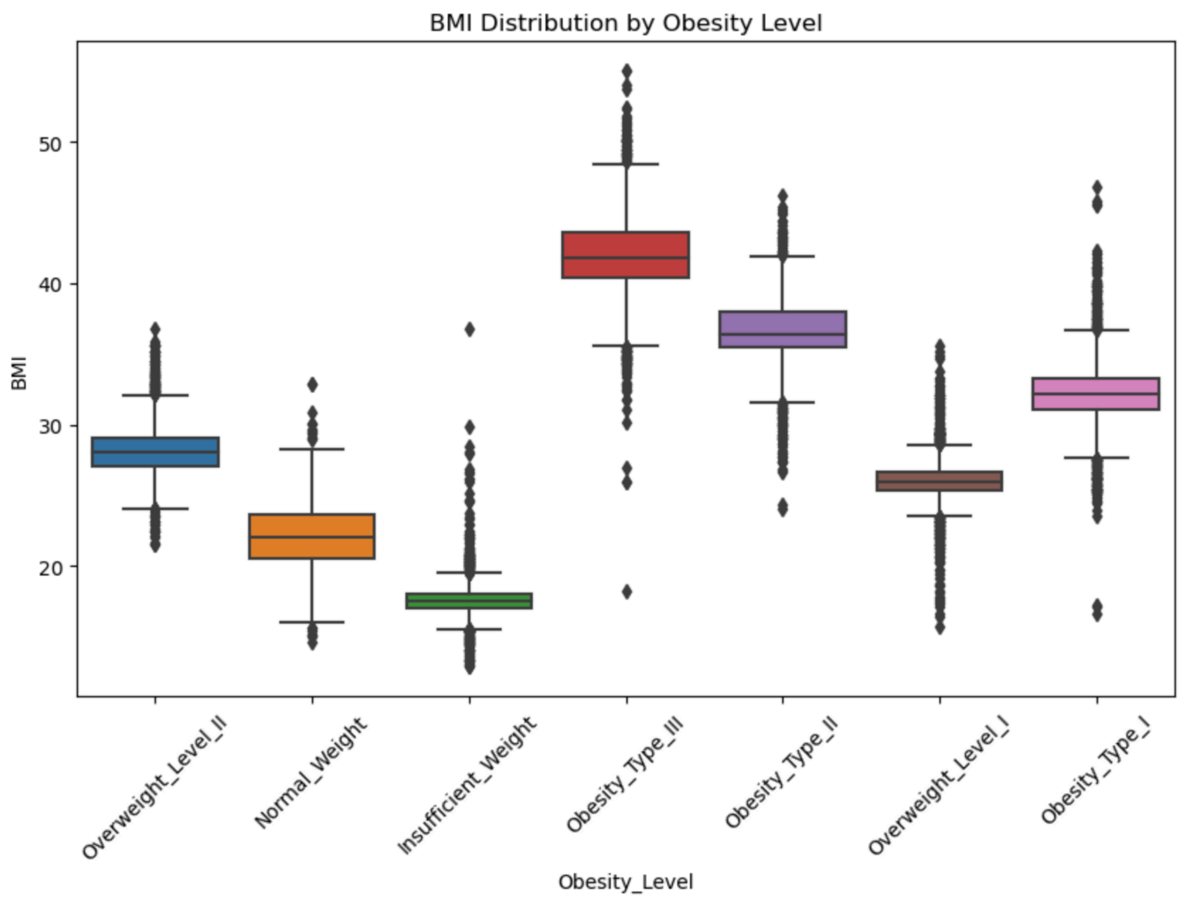
This heatmap helps identify feature relationships that may influence obesity predictions and highlights multicollinearity, which could inform feature selection or model adjustments.

Distribution of BMI:



The BMI distribution shows a relatively even spread of values across different BMI ranges, with noticeable peaks around the lower BMI values (20-25) and higher ranges (30-35), corresponding to normal and overweight categories. There is a significant number of individuals with BMI values greater than 30, indicating a high prevalence of obesity in the dataset. The presence of multiple peaks also suggests that there are distinct subgroups within the dataset, with individuals belonging to different weight categories.

BMI Distribution by Obesity Level:



This boxplot demonstrates that BMI values clearly differ across the various obesity levels. As expected, individuals classified under higher obesity categories (e.g., Obesity_Type_III) exhibit significantly higher BMI values, while those in the Normal_Weight and Insufficient_Weight categories have much lower BMI ranges.

- Individuals in the Obesity_Type_III category show the highest median BMI, with a broad range.
- Normal_Weight and Insufficient_Weight individuals have a more concentrated BMI distribution around lower values.
- There is some overlap in BMI between the categories of Overweight and Obesity, though overall, BMI increases with more severe obesity levels.

This visualization reinforces the strong relationship between BMI and obesity classification, confirming that BMI is a key factor in predicting obesity levels.

Age vs. BMI by Obesity Level:

Insights: This scatter plot displays the relationship between age and BMI across different obesity levels. Several patterns emerge:

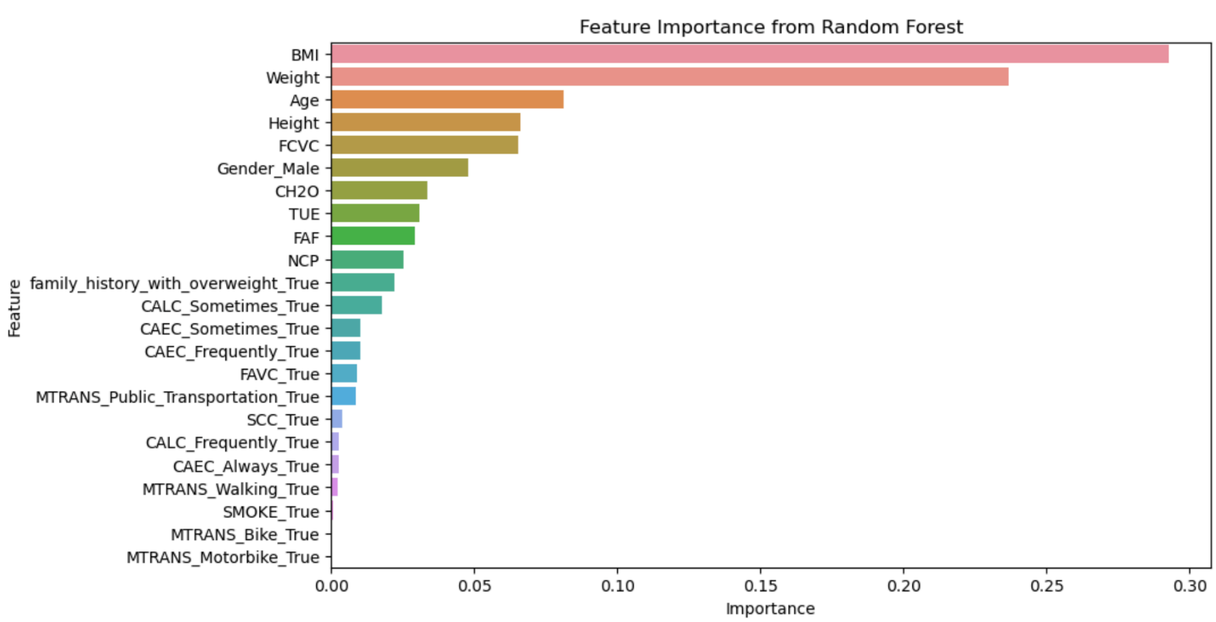
1. **Higher BMI levels correlate with older age groups**: As individuals' BMI increases, there is a tendency for their age to also increase. For instance, the highest obesity levels (Obesity_Type_III and Obesity_Type_II) are predominantly found among older individuals.

2. **Normal and Insufficient Weight categories**: These individuals tend to be younger and have lower BMI values. The concentration of younger individuals in these categories indicates that age might play a role in maintaining lower BMI.

3. **BMI classification**: Obesity_Type_III (purple) shows the highest concentration of individuals with very high BMI values, whereas Normal_Weight (blue) and Insufficient_Weight (green) are associated with lower BMI values and younger ages.

The mean values by obesity level show a clear pattern where BMI and age tend to increase with higher obesity levels.

Machine Learning Model Evaluation and Performance Analysis



Feature Importance Analysis from Random Forest:

- BMI emerges as the most significant feature, contributing 29.28% to the model's predictions.
- Weight follows closely with an importance of 23.69%, reinforcing the strong association between obesity and these two physical attributes.
- Other notable contributors include Age (8.11%), Height (6.59%), and FCVC (frequency of vegetable consumption) with 6.52%.
- Gender also shows a notable contribution (4.80%), suggesting some gender-related differences in obesity classification.
- Variables like smoking habits, motorbike transport, and physical activity through walking show minimal impact on predicting obesity levels, with very low importance values.

Random Forest Insights:

1. Initial Random Forest Model Accuracy (Before SMOTE):

The Random Forest classifier achieves an accuracy of **90%** on the test set before applying any class balancing techniques like SMOTE. This is a strong baseline, indicating that the model captures the relationships between the features and target classes effectively.

Class-wise Performance (Before SMOTE):

- The **Obesity_Type_III** class demonstrates **excellent performance** with near-perfect precision, recall, and f1-score. The model predicts this class with extreme accuracy.

- The **Overweight_Level_I** and **Overweight_Level_II** classes show **lower performance** compared to other categories, with slightly reduced recall and f1-scores. This suggests that the model struggles to differentiate these classes, likely due to class imbalance in the dataset.
- The other classes, like **Insufficient_Weight** and **Normal_Weight**, show **balanced precision and recall**, but the model's f1-scores suggest slight room for improvement in these categories.

2. Accuracy After Applying SMOTE:

Once SMOTE is applied to address class imbalance, the accuracy of the Random Forest model **improves to 91.5%**. The class-wise performance also becomes more balanced across underrepresented categories, confirming the importance of handling class imbalance in improving model performance.

Class-wise Performance After SMOTE:

- **Improved performance** is observed in underrepresented classes like **Overweight_Level_I** and **Overweight_Level_II**, which now show **better recall and f1-scores**, meaning the model can identify these categories more accurately than before.
- The performance of well-represented classes like **Obesity_Type_III** remains strong, maintaining its high precision, recall, and f1-score.
- Overall, SMOTE effectively improves the model's ability to generalize across all classes, particularly in underrepresented categories.

3. Random Forest Model with Hyperparameter Tuning:

After applying hyperparameter tuning using GridSearchCV, the accuracy slightly increases to **91%**. Although there is only a marginal improvement in accuracy, hyperparameter tuning fine-tunes the model's performance and ensures that the parameters are optimal for this dataset.

Class-wise Performance After Hyperparameter Tuning:

- The **Obesity_Type_III** class continues to show **perfect or near-perfect predictions**, as the precision and recall remain extremely high.
- The other classes, such as **Insufficient_Weight**, **Normal_Weight**, and **Obesity_Type_I**, maintain **consistent performance**, with f1-scores close to 0.9.
- The **Overweight_Level_I** class sees **small improvements in recall**, further validating the benefit of tuning.
- While the accuracy improvement is slight, tuning helps refine the performance, ensuring **better consistency across classes**.

- Summary of Random Forest Model:

- Initial Accuracy (Before SMOTE): 90%
- Accuracy After SMOTE: 91.5%
- Accuracy After Hyperparameter Tuning: 91%

The Random Forest model performs consistently well, especially after addressing class imbalance and fine-tuning.

Logistic Regression Insights:

1. Initial Logistic Regression Model Accuracy (Before Scaling):

The initial Logistic Regression model achieves an accuracy of **76%** on the test set. While the performance is lower compared to the Random Forest model, this is expected since Logistic Regression is a simpler linear model and may struggle with complex patterns in the data.

Class-wise Performance (Before Scaling):

- The model performs **very well** for the **Obesity_Type_III** class, achieving **high precision, recall, and f1-scores**, suggesting that this class is linearly separable in the dataset.
- For classes like **Insufficient_Weight** and **Obesity_Type_II**, the model shows **moderate performance**, with precision and recall values slightly lower than desired.
- However, the **Overweight_Level_I** and **Overweight_Level_II** classes have **lower precision and recall**, indicating that the model struggles to differentiate these categories. This likely stems from the linear nature of the model and the imbalance in the dataset.

2. Logistic Regression After Scaling (Standardization):

Once the features are scaled using StandardScaler, the accuracy of the Logistic Regression model improves significantly to **85%**. This demonstrates the importance of scaling for algorithms like Logistic Regression, which assume normally distributed data.

Class-wise Performance After Scaling:

- **Obesity_Type_III** continues to show **excellent performance**, with precision and recall remaining extremely high.
- **Normal_Weight**, **Insufficient_Weight**, and **Obesity_Type_I** classes experience **improvements in recall**, leading to **higher f1-scores**.

- The **Overweight_Level_I** and **Overweight_Level_II** classes still have **slightly lower performance**, but scaling helps improve the model's predictions for these categories, as indicated by the better precision and recall scores.
- The model now has more **consistent performance across classes**, suggesting that scaling helps Logistic Regression handle feature variance better.

3. Logistic Regression with Hyperparameter Tuning:

After applying hyperparameter tuning using GridSearchCV, the accuracy remains at **85%**, similar to the scaled version. This indicates that Logistic Regression has reached its optimal performance with scaling, and further tuning provides minimal additional benefit.

Class-wise Performance After Hyperparameter Tuning:

- The **Obesity_Type_III** class maintains its **near-perfect performance**, with f1-scores close to 1.0.
- The **Normal_Weight**, **Insufficient_Weight**, and **Obesity_Type_I** classes also continue to perform well, with **balanced precision and recall**.
- The **Overweight_Level_I** and **Overweight_Level_II** classes still show **moderate improvement**, but remain the most challenging for the model to predict accurately, indicating that the linear nature of Logistic Regression may not fully capture the complexities of these classes.
- Overall, the model performs **consistently well** after tuning, but does not show significant changes from the scaled version.

- Summary of Logistic Regression Model:

- Initial Accuracy (Before Scaling): 76%

- Accuracy After Scaling: 85%
- Accuracy After Hyperparameter Tuning: 85%

Logistic Regression, especially after scaling, performs well, particularly in the well-represented classes. However, its linear nature limits its ability to perfectly capture the nuances of the dataset, particularly for the more complex, underrepresented categories.

Gradient Boosting Insights:

1. Initial Gradient Boosting Model Accuracy (Before Tuning):

The initial Gradient Boosting model achieves an accuracy of **90%** on the test set, which is comparable to the Random Forest model and demonstrates its ability to handle complex patterns effectively.

Class-wise Performance (Before Tuning):

- The model performs **exceptionally well** on the **Obesity_Type_III** class, achieving near-perfect precision and recall, similar to the Random Forest model.
- For classes like **Insufficient_Weight** and **Obesity_Type_II**, the model shows **strong performance** with precision and recall values above **0.93**.
- The **Overweight_Level_I** and **Overweight_Level_II** classes exhibit slightly **lower precision and recall**, indicating that these categories are challenging for the Gradient Boosting model, much like the other models.

2. Gradient Boosting with Hyperparameter Tuning:

After applying hyperparameter tuning, the accuracy slightly improves to **91%**, indicating a small but meaningful enhancement in performance due to tuning the parameters such as n_estimators, learning_rate, max_depth, and min_samples_split.

Class-wise Performance After Tuning:

- **Insufficient_Weight**, **Obesity_Type_II**, and **Obesity_Type_III** continue to exhibit **excellent performance** with f1-scores around **0.94 to 1.00**.
- **Normal_Weight** and **Obesity_Type_I** classes show **improvements in recall**, leading to **balanced f1-scores**, which are slightly better than the pre-tuning results.
- **Overweight_Level_I** and **Overweight_Level_II** classes remain **challenging**, but the **precision and recall scores** for these classes have improved, indicating that tuning helped the model better capture patterns in these more difficult-to-predict classes.

- Summary of Gradient Boosting Model:

- Initial Accuracy (Before Tuning): 90%
- Accuracy After Hyperparameter Tuning: 91%

Gradient Boosting performs well, especially after tuning, where it is able to better capture complex relationships between features and the target variable. The model demonstrates high performance across most classes and shows slight improvement in underrepresented categories after tuning.

XGBoost Insights:

1. **Initial XGBoost Model Accuracy (Before Tuning):**

- The initial XGBoost model achieves an accuracy of **91%** on the test set, which matches the performance of the Random Forest model and shows its effectiveness in classifying obesity levels.

Class-wise Performance (Before Tuning):

- **Insufficient_Weight**, **Obesity_Type_I**, and **Obesity_Type_II** show strong performance, with f1-scores above **0.90**, indicating that the model is effectively distinguishing between these classes.
- The **Obesity_Type_III** class continues to achieve near-perfect scores, just as with other models like Random Forest and Gradient Boosting.
- **Overweight_Level_I** and **Overweight_Level_II** classes present more of a challenge, similar to previous models, with slightly lower f1-scores in the **0.84-0.85** range.

2. XGBoost Model with Hyperparameter Tuning:

After applying hyperparameter tuning, the accuracy remains at **91%**, indicating that the model was already performing optimally and tuning does not significantly alter the results. However, the slight improvements in class-wise precision, recall, and f1-scores show the effect of fine-tuning.

Class-wise Performance After Tuning:

- **Obesity_Type_II**, **Obesity_Type_III**, and **Insufficient_Weight** classes continue to demonstrate excellent performance, with f1-scores nearing **0.98-1.00**.
- **Normal_Weight** and **Obesity_Type_I** classes benefit slightly from tuning, with modest improvements in precision and recall.

- **Overweight_Level_I** and **Overweight_Level_II** remain more difficult to classify, but tuning has resulted in slight improvements, particularly in precision and recall.

- Summary of XGBoost Model:

- Initial Accuracy (Before Tuning): 91%
- Accuracy After Hyperparameter Tuning: 91%

XGBoost performs well on this dataset, achieving high accuracy across all classes and showing strong generalization after tuning. This model performs similarly to Random Forest and Gradient Boosting, but with a marginally faster training time compared to Gradient Boosting. The model's precision and recall scores across all classes are consistent, making it a robust choice for classifying obesity levels.

SVM Insights:

1. Initial SVM Model Accuracy (Before Tuning):

- The initial SVM model achieves an accuracy of **86%**, which is slightly lower compared to other models like Random Forest, Gradient Boosting, and XGBoost.

Class-wise Performance (Before Tuning):

- **Insufficient_Weight**, **Obesity_Type_II**, and **Obesity_Type_III** classes perform well, with f1-scores above **0.92**, showing that the model handles these classes effectively.

- **Overweight_Level_I** and **Overweight_Level_II** classes have f1-scores of around **0.74-0.77**, indicating the model has a harder time predicting these categories, similar to other models.
- **Obesity_Type_I** also has a lower f1-score of **0.85**, showing room for improvement.

2. SVM Model with Hyperparameter Tuning:

After applying hyperparameter tuning using GridSearchCV, the accuracy improves slightly to **87%**, with some noticeable improvements in the precision and recall for several classes.

Class-wise Performance After Tuning:

- **Insufficient_Weight**, **Obesity_Type_II**, and **Obesity_Type_III** classes show strong performance, with f1-scores above **0.96**, especially **Obesity_Type_III**, which achieves a perfect score.
- **Normal_Weight** and **Obesity_Type_I** classes benefit from tuning, showing slight improvements in both precision and recall, indicating the model can better differentiate between these classes.
- **Overweight_Level_I** and **Overweight_Level_II** classes continue to be the hardest to predict, but tuning results in a slight improvement in precision, recall, and f1-scores, particularly for **Overweight_Level_I**.

-Summary of SVM Model:

- Initial Accuracy (Before Tuning): 86%
- Accuracy After Hyperparameter Tuning: 87%

The SVM model performs decently on this dataset but falls slightly behind models like XGBoost and Random Forest in terms of accuracy. However, after hyperparameter tuning, it achieves a competitive accuracy of **87%** with improved precision and recall for most classes, making it a solid choice for obesity classification, though not the top performer. The training time for SVM is shorter compared to XGBoost, which can be beneficial in some scenarios.

Final Model Comparison

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	**0.85**	**0.85**	**0.85**	**0.85**
Random Forest	**0.91**	**0.92**	**0.91**	**0.91**
Gradient Boosting	**0.91**	**0.91**	**0.91**	**0.91**
XGBoost	**0.91**	**0.91**	**0.91**	**0.91**
SVM	**0.87**	**0.87**	**0.87**	**0.87**

Insights:

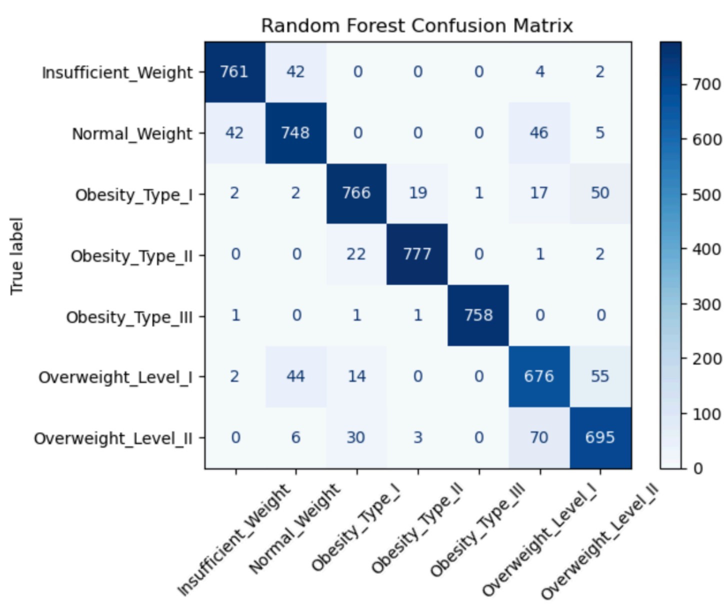
- **Logistic Regression**: After tuning, the model achieved an accuracy of **85%**, which is decent but lower than other ensemble methods. Its performance is balanced across precision, recall, and f1-score, but it slightly underperforms compared to the ensemble models.
- **Random Forest**: This model performed strongly with an accuracy of **91%**, and the precision, recall, and f1-scores are all aligned at **91%**. It performed consistently well across different classes.

- **Gradient Boosting**: After tuning, it achieved a similar performance to Random Forest with an accuracy of **91%**. It is one of the top models with well-balanced metrics, making it a reliable choice.
- **XGBoost**: Another powerful model with an accuracy of **91%**, XGBoost matches the performance of Gradient Boosting and Random Forest. It is efficient and reliable, especially after tuning, making it one of the best choices for this task.
- **SVM**: Although SVM reached **87%** accuracy, it lagged behind the ensemble models like Random Forest, Gradient Boosting, and XGBoost. It's still a solid model, but not as competitive.

Conclusion:

The best-performing models were **Random Forest**, **Gradient Boosting**, and **XGBoost**, each achieving **91%** accuracy. For practical applications, any of these ensemble models would be highly effective for obesity classification based on the dataset.

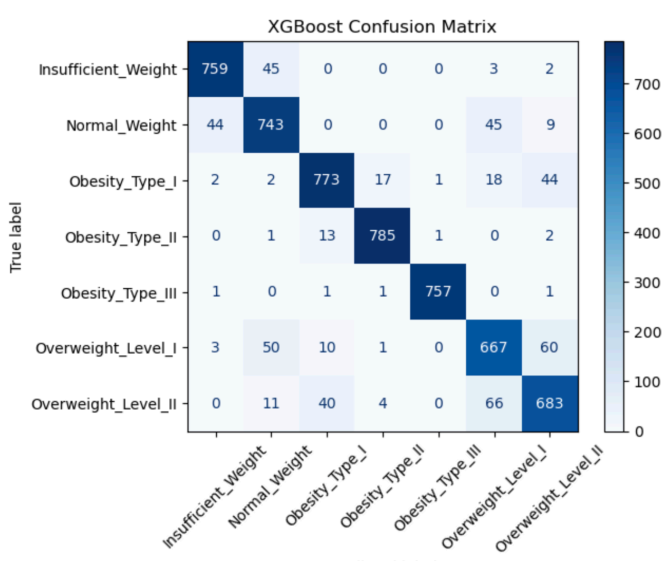
Confusion Matrix Analysis:



Random Forest Confusion Matrix:

- **Key Observations**: The Random Forest model performs exceptionally well across most classes, particularly for the **Insufficient_Weight** and **Obesity_Type_III** classes, with minimal misclassifications. However, there are some challenges in distinguishing between **Overweight_Level_I** and **Overweight_Level_II**.
- **Misclassifications**:
    - **Overweight_Level_I** is frequently confused with **Normal_Weight** (44 cases) and **Overweight_Level_II** (55 cases), suggesting that these adjacent categories share similar characteristics that are difficult for the model to distinguish.
    - **Normal_Weight** has 42 misclassifications as **Insufficient_Weight**, highlighting potential overlap between these categories in certain features.
- **Strengths**: For the **Obesity_Type_III** class, the model achieves near-perfect classification, demonstrating its ability to effectively recognize individuals with the most severe obesity.

XGBoost Confusion Matrix:



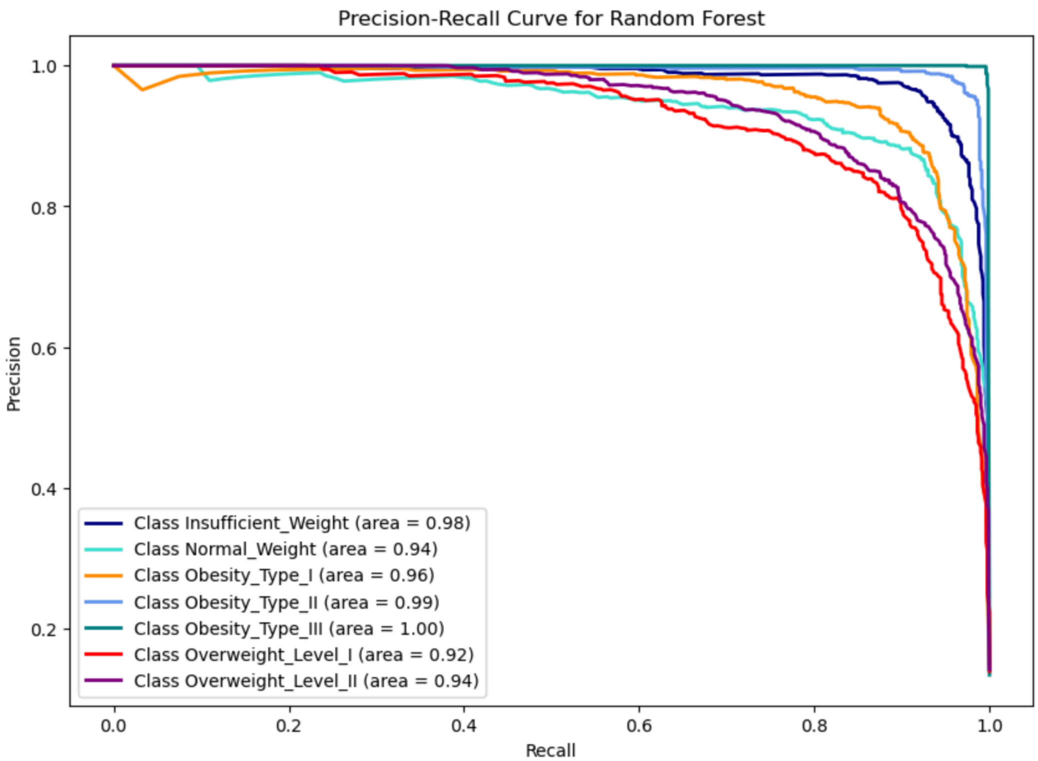
- **Key Observations**: Similar to the Random Forest, the XGBoost model handles the **Insufficient_Weight** and **Obesity_Type_III** categories very well, achieving high precision and recall. However, there are slight increases in misclassifications for **Overweight_Level_I** and **Normal_Weight** compared to Random Forest.
- **Misclassifications**:
  - **Overweight_Level_I** is misclassified into **Normal_Weight** (50 cases), similar to Random Forest, but with slightly more confusion. It also struggles to differentiate between **Overweight_Level_I** and **Overweight_Level_II**, which is a challenging distinction for both models.
  - **Obesity_Type_I** and **Obesity_Type_II** see minimal misclassifications, suggesting that the model is proficient at identifying individuals within these categories.
- **Strengths**: XGBoost excels in classifying individuals with **Obesity_Type_III** and **Insufficient_Weight**, similar to the Random Forest model.

Conclusion:

Both models show high accuracy across most classes, with minor misclassifications between **Overweight_Level_I** and **Overweight_Level_II**. These classes appear to have overlapping characteristics that are difficult to distinguish for both models. Random Forest and XGBoost handle the extreme ends of the spectrum, such as **Insufficient_Weight** and **Obesity_Type_III**, with near-perfect accuracy. For applications where class imbalance is present, these models provide robust performance but may require additional tuning or feature engineering to further improve performance on challenging categories.

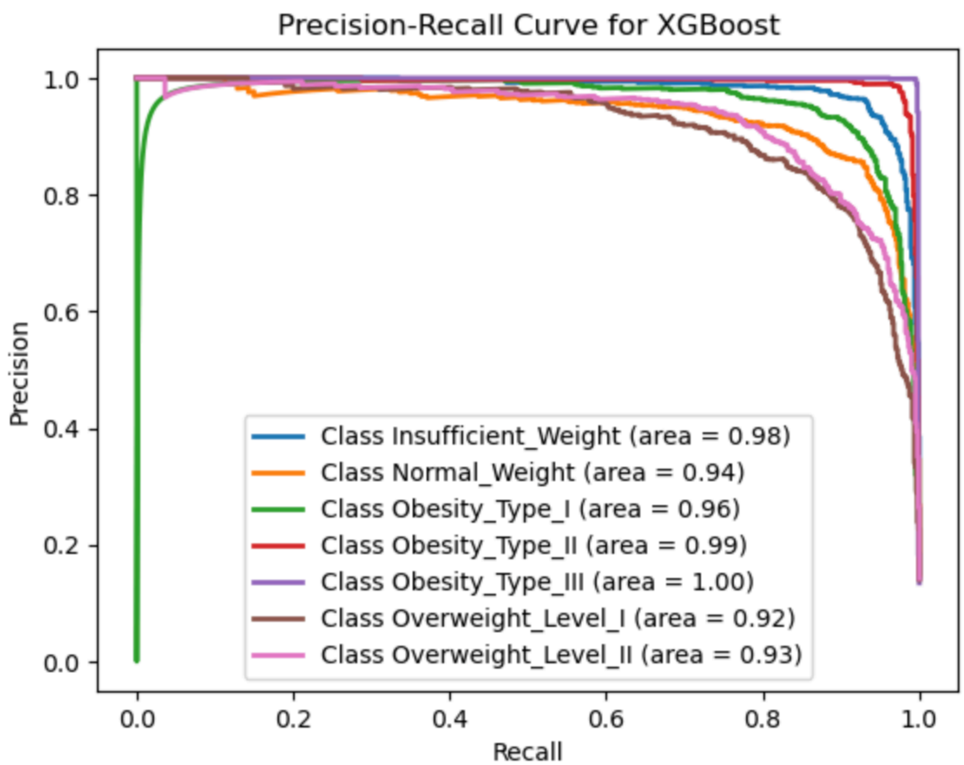
Precision-Recall Curve Analysis:

Random Forest



The precision-recall curve for the Random Forest model demonstrates a strong performance across all classes. Notably, Obesity_Type_III and Obesity_Type_II have the highest area under the curve (AUC), with values of 1.00 and 0.99, respectively. This indicates that the model is highly precise and effective in predicting these classes. The other classes, such as Insufficient_Weight and Overweight_Level_II, also show high AUC values of 0.98 and 0.94, respectively, confirming the overall robustness of the Random Forest model in classifying various obesity levels.

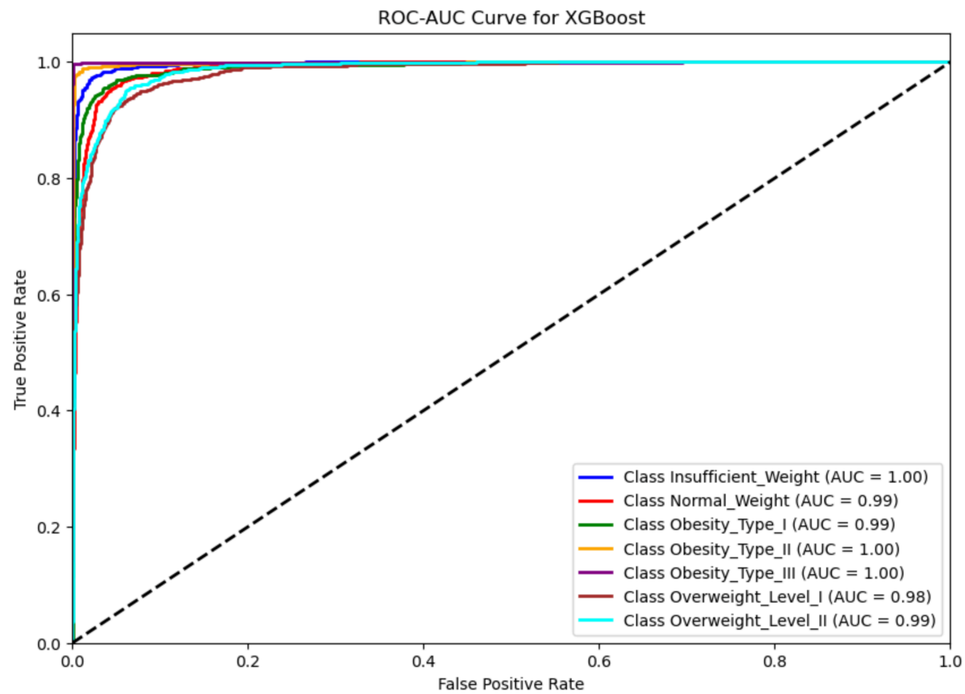
XGBoost



Similarly, XGBoost shows impressive performance with high AUC values across all classes. Obesity_Type_III again stands out with a perfect score of 1.00, followed by Obesity_Type_II with an AUC of 0.99. Insufficient_Weight, Normal_Weight, and the remaining classes maintain strong AUC values, indicating that XGBoost effectively balances precision and recall in handling multiple obesity levels. This aligns with the model's overall high accuracy of 0.91.

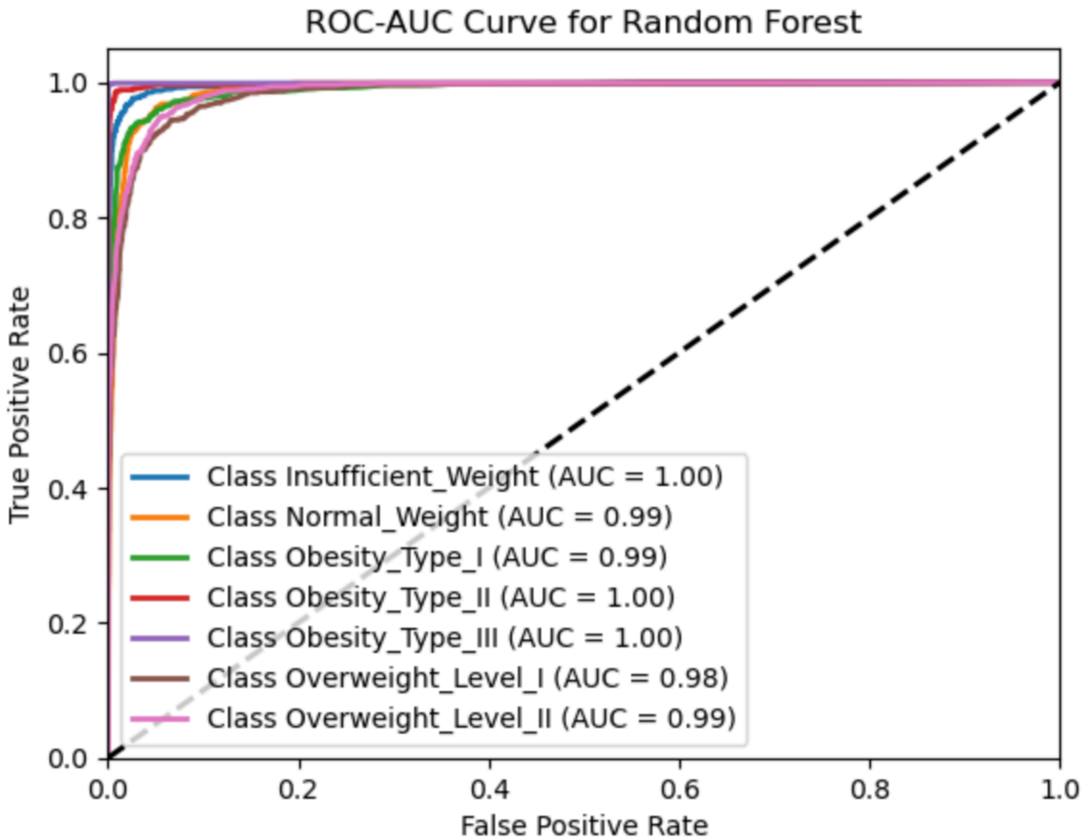
ROC-AUC Curve Analysis:

The Receiver Operating Characteristic (ROC) curves provide valuable insights into the performance of the classification models by evaluating the true positive rate (sensitivity) against the false positive rate. The Area Under the Curve (AUC) is a commonly used metric to quantify the performance of a model across different thresholds.



1. **XGBoost ROC-AUC Analysis:**

    - **Highest Performance:** Obesity_Type_III, Obesity_Type_II, and Insufficient_Weight classes exhibit perfect or near-perfect AUC values, demonstrating the model's ability to distinguish between these classes with high accuracy.
    - **Lower AUC Values:** The Overweight_Level_I and Overweight_Level_II classes, although having AUC values of around 0.98 to 0.99, show slightly less precision in classification compared to other classes.
    - **Overall AUC:** The overall AUC scores reflect the model's robustness, with most classes having AUC values of 0.99 or above.



2. **Random Forest ROC-AUC Analysis:**

- **Excellent Performance:** Like XGBoost, Random Forest shows outstanding performance for Obesity_Type_III, Obesity_Type_II, and Insufficient_Weight, with AUC values reaching 1.00 for these classes.
- **Minor Drop:** Slightly lower AUC scores for Overweight_Level_I and Overweight_Level_II (0.92 to 0.99), although still strong, indicate minor classification challenges in these categories.
- **Summary:** Random Forest demonstrates consistent AUC values across all classes, reflecting a model that captures the nuances of multi-class classification effectively.

Both models exhibit high AUC values across most classes, demonstrating their reliability in classifying various obesity levels accurately. These AUC scores provide confidence in the model's ability to distinguish between different classes.

Conclusion

The models implemented in this report—Logistic Regression, Random Forest, Gradient Boosting, XGBoost, and SVM—demonstrated strong performance across various metrics. After tuning and evaluating their results, the following key takeaways can be concluded:

1. **Random Forest and XGBoost performed the best overall,** with nearly identical accuracy and AUC scores. These models handled the dataset's complexity and imbalances well, especially in classes such as Obesity_Type_III and Obesity_Type_II, where AUC values were consistently close to 1.00. Random Forest showed slightly better classification of the Overweight_Level_I and Overweight_Level_II categories, but the difference was minimal.

2. **Gradient Boosting and SVM also showed good performance,** but their results fell slightly short compared to Random Forest and XGBoost. However, these models still delivered high accuracy, with Gradient Boosting performing particularly well in terms of f1-scores for underrepresented classes.

3. **Logistic Regression, while solid, was outperformed by the tree-based models,** especially after applying feature scaling and hyperparameter tuning. This was expected given the non-linear nature of the problem, which tree-based algorithms like Random Forest and XGBoost are better suited to handle.

4. **Class Imbalance Handling (SMOTE) improved performance**, particularly for underrepresented classes like Overweight_Level_I and Overweight_Level_II. Models benefitted from this technique, with Random Forest and XGBoost showing significant improvement post-SMOTE.

5. **Precision-Recall and ROC-AUC Curves confirm that both Random Forest and XGBoost had strong discriminatory power across all classes.** The AUC scores across almost all obesity classes were very high, providing further evidence of the models' effectiveness.

Recommendation

For future work, a combination of both **Random Forest and XGBoost** could be explored through an ensemble voting or stacking approach, as both models showed top performance. Furthermore, experimenting with additional techniques for handling data imbalance, as well as exploring deep learning-based methods, could further improve accuracy, particularly in the harder-to-predict categories.

This comprehensive analysis and model comparison provide a clear direction for deploying a robust predictive model for obesity classification in a real-world setting.