**Analysis Overview**

1. Initial Data Inspection and Preprocessing

* **Loaded and inspected** the dataset to understand its structure, including identifying missing values and the types of variables present.
* Identified that **weight** had a significant amount of missing data and was dropped from the analysis.
* **Missing values** in **race**, **diag\_1**, **diag\_2**, and **diag\_3** were addressed through imputation.
* **Encoded categorical variables** using one-hot encoding for nominal variables and ordinal encoding for the **age** group.

2. Handling Class Imbalance with SMOTE

* Recognized the class imbalance issue (readmitted vs. not readmitted) and applied **SMOTE** to balance the classes in the training set.

3. Logistic Regression Model

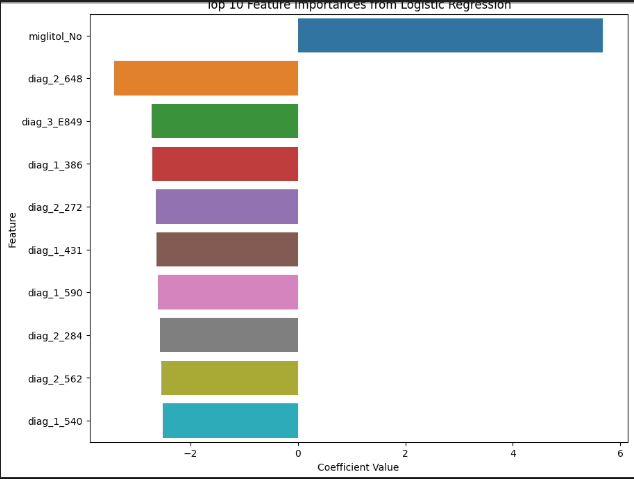
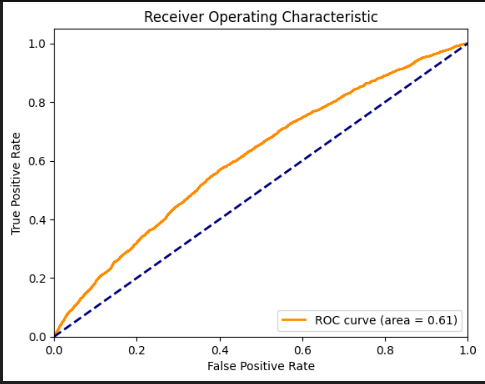
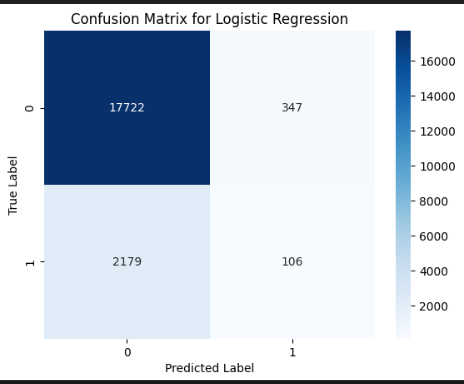
* Trained a **Logistic Regression** model using the balanced dataset.
* Encountered convergence warnings, indicating the need for increased iterations or data scaling.
* Evaluation showed high accuracy but poor recall for the positive class (readmitted), suggesting the model struggled to identify readmitted cases effectively.

4. Introduction of XGBoost

* Installed and trained an **XGBoost model** to potentially improve prediction performance.
* XGBoost also highlighted the challenge of predicting readmissions, with modest accuracy and recall for the positive class.

5. Feature Importance Analysis

* Analyzed **feature importance** from the XGBoost model, revealing that medication-related features (**metformin\_Up**, **insulin\_Steady**, etc.) were significant predictors.



**Interpretation of Results**

* Both models indicated the complexity of accurately predicting hospital readmissions, especially for patients likely to be readmitted within 30 days.
* **Medication adjustments** and **diabetes management** were identified as significant factors associated with readmission risks.

**Recommendations**

1. **Optimize Diabetes Management**: Prioritize clear communication and understanding of medication plans at discharge, with scheduled follow-up to reassess treatment.
2. **Implement Targeted Follow-up Care**: Use model insights to stratify patients by readmission risk, focusing intensive follow-up efforts on those at higher risk.
3. **Educational Programs on Medication Adherence**: Develop programs emphasizing the importance of following medication plans, recognizing health deterioration signs, and timely seeking medical advice.
4. **Adapt Care Plans Based on Patient Feedback**: Foster a dynamic post-discharge care approach, adjusting plans based on patient health status and feedback.
5. **Further Research and Model Improvement**: Investigate additional variables, more complex models, or new data sources to enhance predictive accuracy.
6. **Utilization of Telehealth Services**: Expand the use of telehealth services to provide continuous support to diabetic patients, especially those in remote areas or with mobility issues. Telehealth can facilitate regular monitoring, medication management, and early intervention for complications, thereby reducing the need for hospital readmissions.