Model Training and Hyperparameter Tuning:
The dataset was split into a training set and a test set using a 90-10 split, ensuring stratification to maintain a similar proportion of positive labels in both sets. Features were standardized using StandardScaler to normalize the data, essential for effective model training, especially for models like Logistic Regression that are sensitive to variable scales. Multiple models were trained, including Logistic Regression, Random Forest, and Gradient Boosting. Hyperparameters for each model were tuned using GridSearchCV with a 5-fold cross-validation to find the optimal settings that maximize accuracy. The training process was comprehensive, ensuring that the models could generalize well over unseen data.

Logistic Regression:

Emerged as the superior model with an optimal regularization parameter C of 0.01.
Achieved the highest cross-validation score of 0.817.
Test set performance yielded an accuracy of 73.3%.
Notable precision of 75% for predicting the more critical class, underscoring its efficacy in medical diagnostic settings.
Gradient Boosting:

Incorporated interaction terms like age times trestbps (ageXtrestbps) and smoke times oldpeak (smokeXoldpeak).
Exhibited an accuracy of 75.6%, enhancing the predictive accuracy seen in Logistic Regression.
Highlighted by balanced precision and recall across both classes.
Fine-tuned with a learning rate of 0.1, depth of 3, and 100 estimators, achieving consistent performance under varying data conditions.

Comparing:
Involved Logistic Regression, Random Forest, and Gradient Boosting without additional data transformations or interaction terms.
Gradient Boosting consistently proved to be the best model, mirroring the previous accuracy and demonstrating robustness without data transformations.
Final script included specific interaction terms and transformations, confirming the robustness of Gradient Boosting under varying data conditions.
Conclusion:
The best model result achieved was model_5, the gradient boosting model with parameters {'learning_rate': 0.1, 'max_depth': 3, 'n_estimators': 100}, which achieved a cross-validation score of 0.805 and an accuracy of 75.6% on the test set. The confusion matrix shows 30 true negatives and 38 true positives, indicating effective classification. Precision for detecting the more prevalent class is high at 79%, with a recall of 76%, leading to a balanced f1-score of 78%. Overall, the model demonstrates robust performance with a good balance between precision and recall across classes. The Logistic Regression model follows closely in second place.