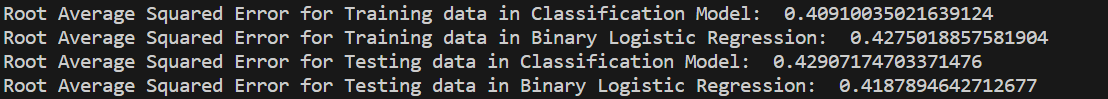
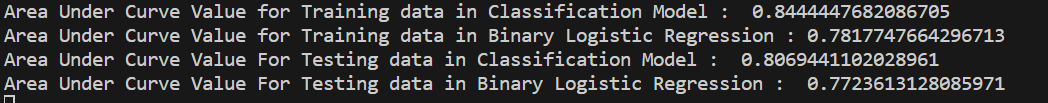
Question 1)

Root Average Squared Errors are:



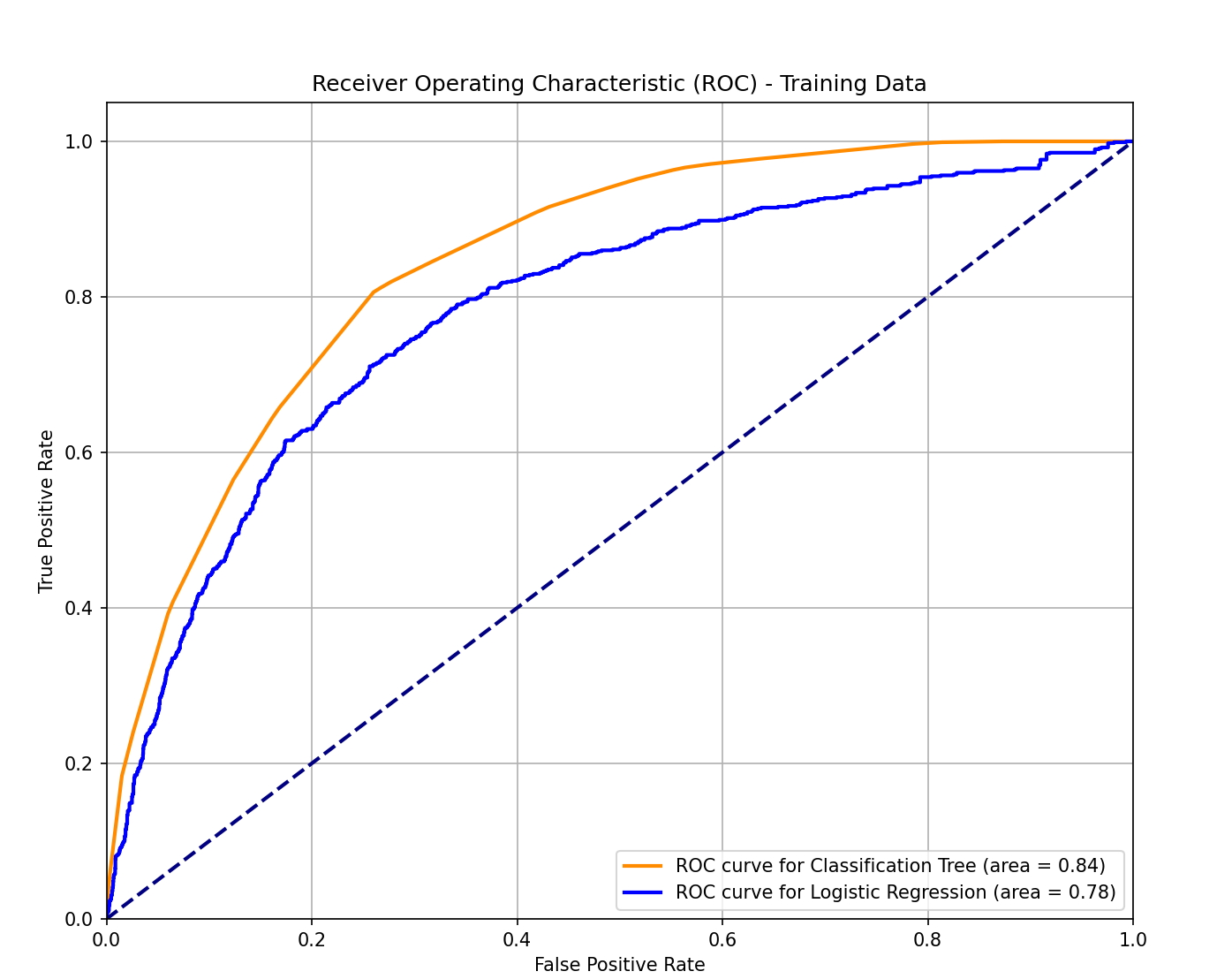
Question 2)

AUC Values are:



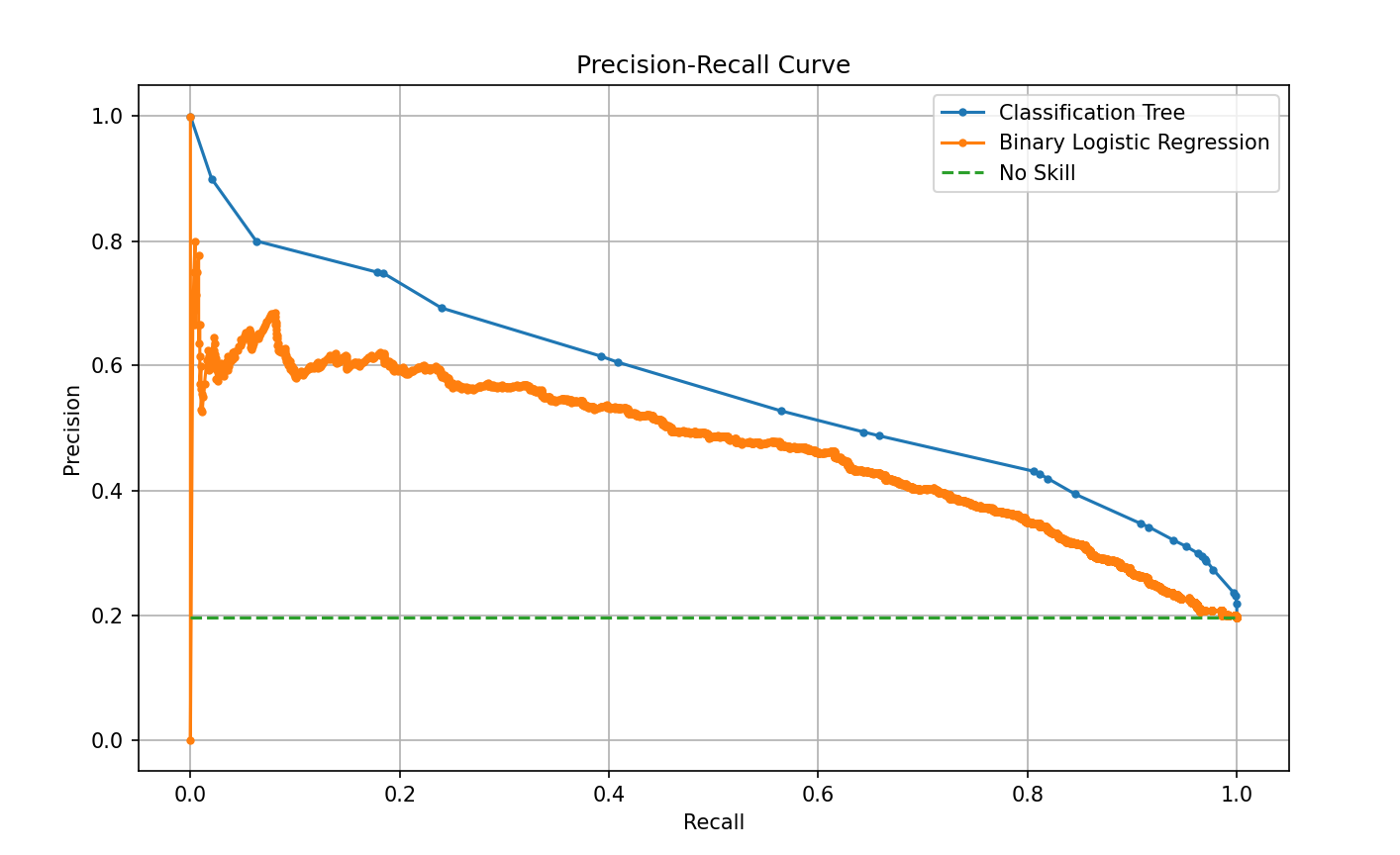
Question 3)

After plotting the ROC Curves:



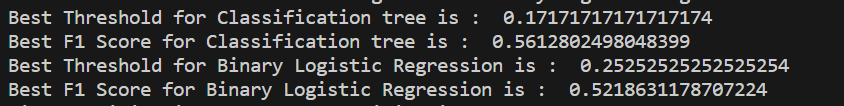
Question 4)

After plotting Precision Recall Curve:



Question 5)

Best Thresholds and F1-Scores for both the models are:



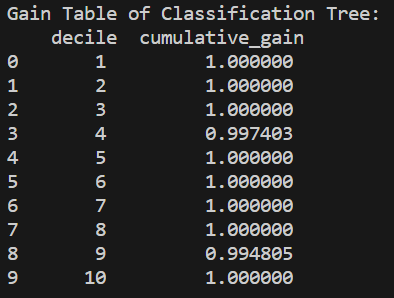
Question 6)

F1-based Misclassification rates for both the models are:

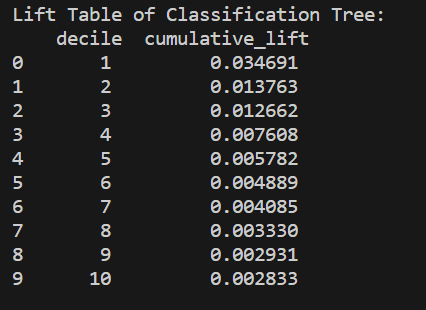


Question 7)

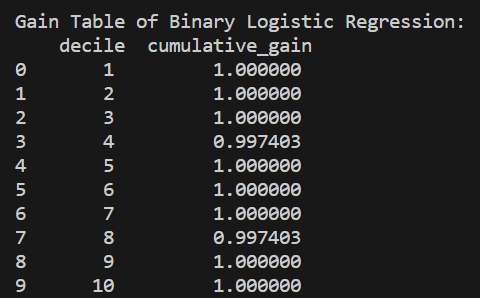
* Cumulative Gain for Classification Model:



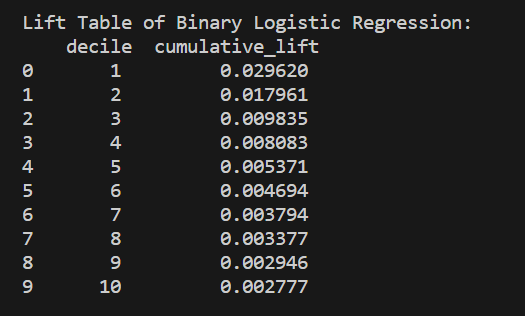
* Cumulative Lift for Classification Model:



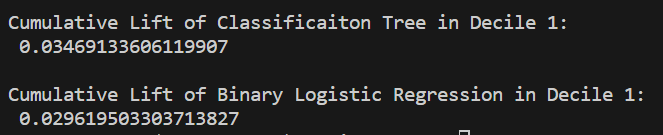
* Cumulative Gain for Binary Logistic Regression:



* Cumulative Lift for Binary Logistic Regression:



* ‘Classification Tree’ has the highest Lift value in Decile 1:



Question 8)

* Root Average Squared Error (RASE):
* Training Data:
* Classification Model: 0.4091
* Binary Logistic Regression: 0.4275
* Testing Data:
* Classification Model: 0.4291
* Binary Logistic Regression: 0.4188
* Area Under Curve (AUC):
* Training Data:
* Classification Model: 0.8444
* Binary Logistic Regression: 0.7818
* Testing Data:
* Classification Model: 0.8069
* Binary Logistic Regression: 0.7724
* F1 Score:
* Classification Model:
* Best Threshold: 0.1717
* Best F1 Score: 0.5613
* Binary Logistic Regression:
* Best Threshold: 0.2525
* Best F1 Score: 0.5219
* Misclassification Rates (based on F1 Score):
* Classification Model: 0.2513
* Binary Logistic Regression: 0.2328
* Cumulative Lift (Decile 1):
* Classification Model: 0.0347
* Binary Logistic Regression: 0.0296

Based on these metrics, out of 5 performance metrics ‘Classification Tree’ performs better in 3, so Classification Tree would be better than Binary Logistic Regression:

* Based on F1 Score, Classification model performs slightly better.
* If minimizing misclassification rates is crucial, you might favor the model with the lower F1-based misclassification rate. In this case, Binary Logistic Regression performs slightly better.
* If maximizing AUC is a priority, the Classification Model might be preferred.
* Considering lift value, the Classification Model has a higher cumulative lift in Decile 1.
* Based on RASE value, Binary Logistic Regression performs better.