Linear model

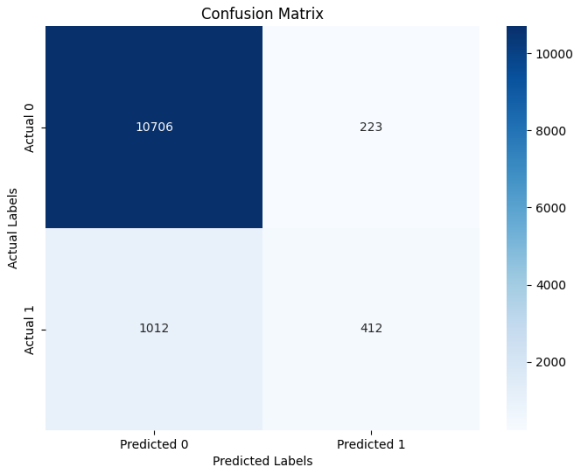
1. Classification Report

category	precision	recall	F1-score
0	0.91	0.98	0.95
1	0.65	0.29	0.40

	precision	recall	F1-score
Accuracy			0.90
Macro avg	0.78	0.63	0.67
Weighted avg	0.88	0.90	0.88

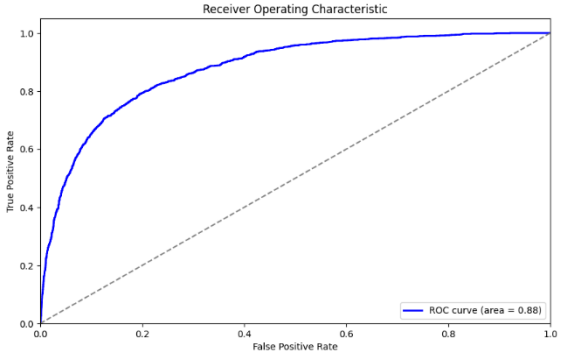
The results from the two tables demonstrate that this model is effective in classifying category 1, indicating that the majority of customers are unlikely to purchase the bank's product service. However, the performance in category 1 is not as strong as desired. One contributing factor could be the imbalanced sample and limited sample size.

2. Confusion matrix



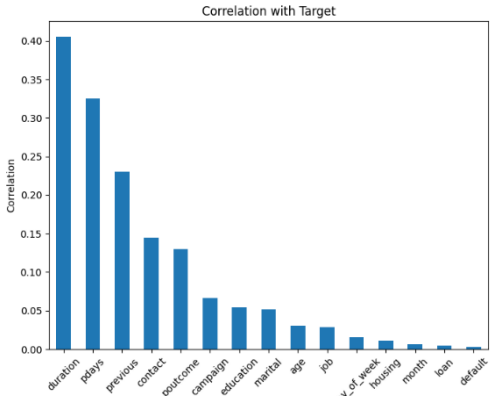
The overall accuracy of the model is approximately 90%. The majority of the samples are predicted to be category 0. However, the model's performance could be improved.

3. ROC curve



The AUC value obtained by this model is 0.8828, indicating that the model has high accuracy and stability in distinguishing positive and negative samples. This is an excellent performance, and the AUC value also shows that the linear model has high robustness and can perform well with a low false alarm rate.

4. Correlation between independent and dependent variables



The plot indicates that the duration of the last contact has the strongest correlation with the dependent variable, suggesting that this factor has a significant impact on the likelihood of a customer purchasing the bank's products. Additionally, the number of days that have elapsed since the last contact also has a notable influence on the dependent variable. However, the presence of a default credit status or loan does not appear to have a substantial effect on the dependent variable.

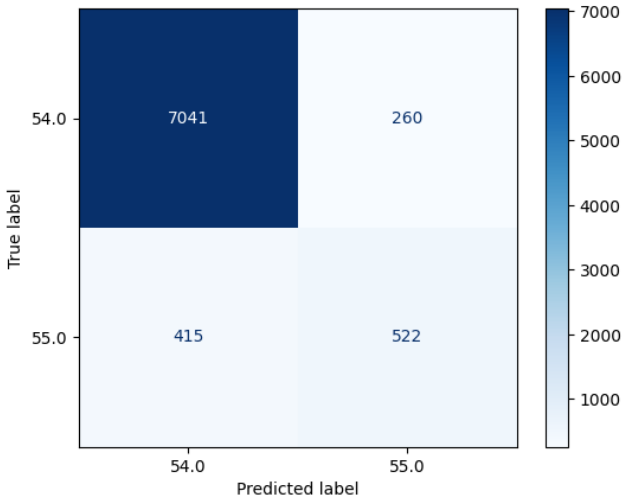
5. Compare with other models:

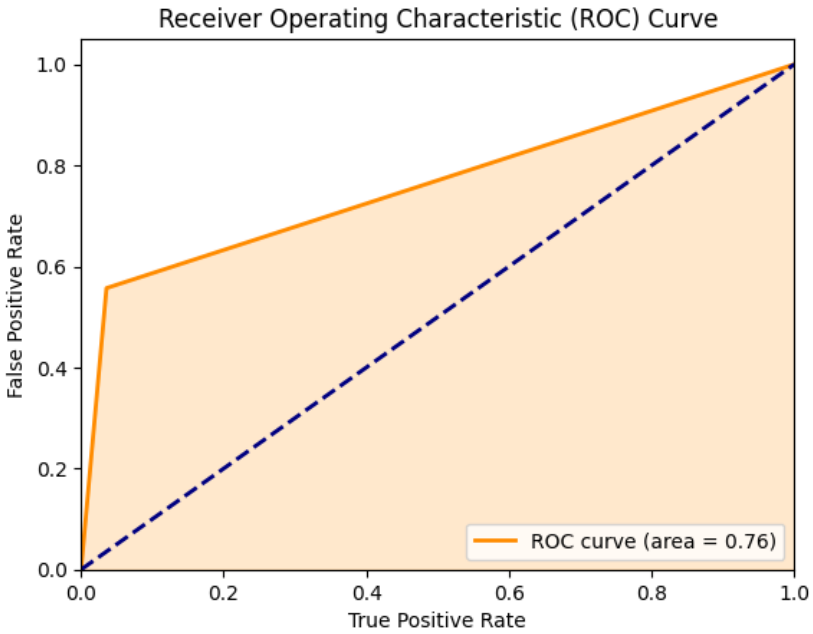
   a. Ensemble (AUC: around 0.67)

category	precision	recall	F1-score
0	0.92	0.97	0.95
1	0.65	0.37	0.47

	precision	recall	F1-score
Accuracy			0.90
Macro avg	0.78	0.67	0.71
Weighted avg	0.89	0.90	0.89

   b. Boosting (Acc.: 0.91%)





The results demonstrate that the boosting model exhibits slight superiority over the linear model in terms of overall performance. The ensemble model, on the other hand, demonstrates comparable performance to the linear model, with the notable distinction of more accurately predicting category 1. However, both models' AUC value is lower than that of the linear model, indicating that they may not be as robust as the latter. Furthermore, since the linear model is the simplest of the three, it can be calculated more rapidly than the other two, resulting in greater efficiency than the boosting and ensemble models.