Use Case: Reduce the marketing cost by X% and acquire Y% of the prospects (compared to random calling), where X and Y are to be maximized

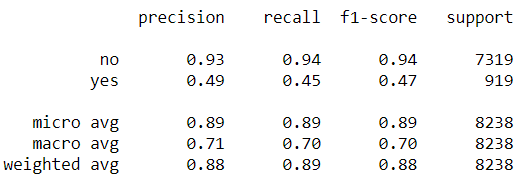
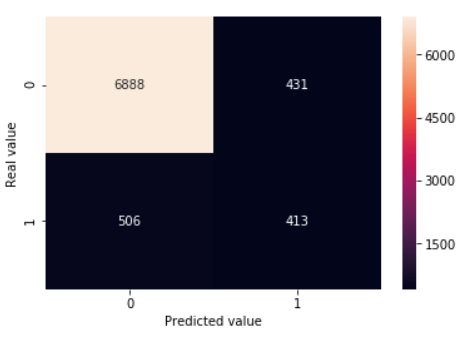
Machine Learning Model: Gradient Boosting Classifier

Preprocessing:

1. Oversampling (SMOTE) because this is an unbalanced dataset.

Training accuracy: 88.67%

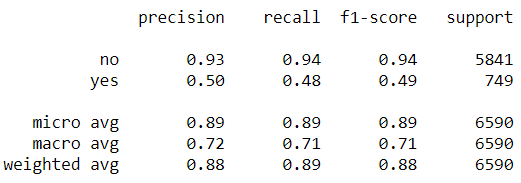
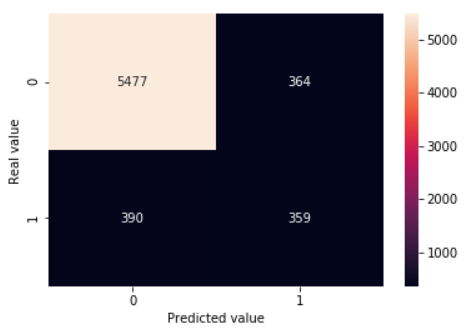
Testing accuracy : 88.62%



Okay, so this model is performing well with test accuracy of 88%. We have a high recall (true positive rate) for target class “no” with 94%. Whereas for target class “yes” it is 45%, which increased from 22% in my baseline model.

Using over-sampling technique to increase the samples for minority class, we tend to increase the recall percentage of customers who actually subscribed to a term deposit.

Validation accuracy : 88.5%



The model is performing even better for the validation set, giving us recall % of 48% for target class “yes”, while maintaining the accuracy for the target class “no”.

How will this help?

Let’s look at the results of the validation set.

Out of 749 customers who actually subscribed to a term deposit, it predicted correctly 48% (359) of the time.

Also, out of 5841 customers who actually did not subscribe, it predicted incorrectly 6% (364) of the time, i.e. false positive.

Let’s say, we have 6,590 customers, and if you are to call each one of them, say you spend rs. 1 per call, that would make total cost at rs. 6590. And say only 10% of them turned, so keeping a gain of rs.10 for all customers who subscribed, the total earnings will be 659 \* 10 = 6590 rs. So when we subtract total loss from the total earnings, we are left with no money.

Now according to the model, it predicted for 723 (359 + 364) customers who will subscribe to a long term deposit. So, total cost to call these customers will be 723 rs. Since the model predicted that 48% of them actually subscribed, so our total earnings will be 347 \* 10 = 3470. Now when we subtract total cost from total earnings, we get 2747 rs.

So, using machine learning, we are able to predict the customers who will subscribe to term deposit and make money.