Non - Owner Class (Action Taken) Prediction

As we know, we faced an issue to predict properly the non-owner class (Action taken) as we had imbalanced target variables like 3 % Buy and 97 % NA.   
But after doing following steps, we are achieving satisfactory results:

Step1: Import 12 QRT data and removed unnecessary columns (like Fund\_id,Stock\_Id,status,ClassDate) from the dataset.

Step2: Normalization only on Numeric data and other data preprocessing.

Step3: Split data based on last qtr date (6/30/2018)

Step4: Again, split data based on second last qtr date (3/31/2018)

Step5: Train model on last 10 QTR data. we have used XBGBoosting Model for Classification. The two reasons to use XGBoost are also the three goals of the project:

* Execution Speed.
* Model Performance.
* Feasibility — easy to tune parameters and modify objectives
* Before train model, we tuned Scale\_pos\_weight. Generally, the Scale\_pos\_weight is the ratio of number of negative class to the positive class. Suppose, the dataset has 90 observations of negative class and 10 observations of positive class, then ideal value of scale\_pos\_Weight should be 9.

Step6: Apply trained model on last QTR data , and we are getting classification matrix as bellow:

NA Buy

NA 59168 115

Buy 2814 2646

Step7: Validation Measure:   
 Error rate : 0.04524041209088241

Accuracy : 0.9547595879091176

Sensitivity : 0.4846153846153846

Specificity : 0.4846153846153846

False positive rate : 0.0019398478484557123