Great Learning & UT Austin

Prediction whether the customer is going to adopt the tourism package based on a social media campaign.

Social Media_Tourism_ Project

Project Notes-2

Submitted By- Gunjar Fuley
Batch- PGPDSBA Online Nov_A 2020
Email- gforgunjaar@gmail.com
Phone- 9938126651

GO-GO AIR

Effective Social Media Campaign for higher revenue through an increase in sales of tickets

1. MODEL BUILDING AND INTERPRETATION

In the project notes 1, we have understood various inconsistencies with the dataset like the missing values, outliers etc. Based on the understanding the treatment was done on the dataset like the missing value treatment, outlier treatment, labelling the data etc. The dataset was further divided into two parts based on Login devices i.e. either Mobile or Laptop. Subsequently the feature engineering was applied and 4 features were removed for better results.

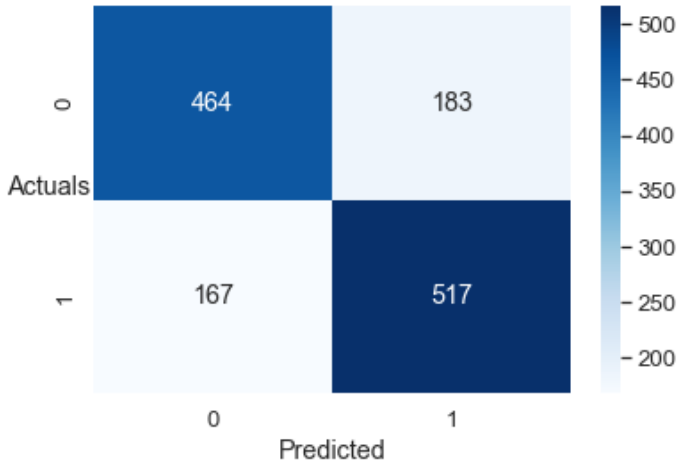
Before building the model the data was scaled using the Standard Scalar. Also, the dataset was split into train & test datasets in the ratio of 80:20. 80% is the Train Set and 20% is the Test Set.

a. Building various models

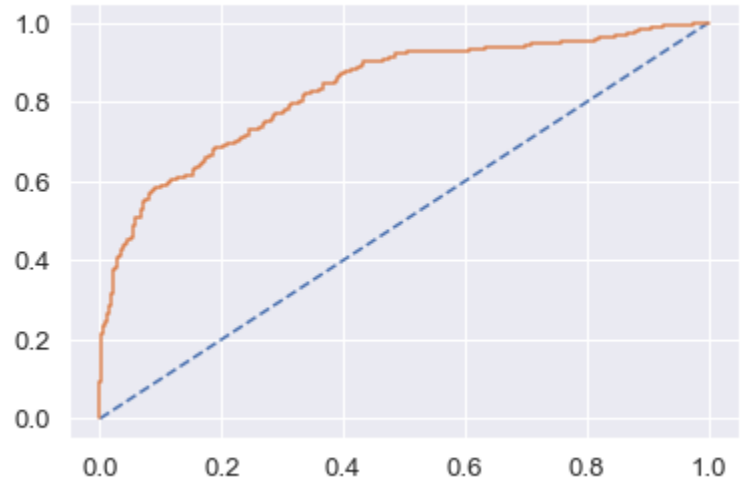
LAPTOP-

Laptop Logistic Regression Training Set

```
The model score for Logistic Regression Training set is 0.7370398196844478

The classification report & Confution matrix for Logistic Regression training set is
              precision    recall  f1-score   support

         0.0       0.74      0.72      0.73       647
         1.0       0.74      0.76      0.75       684

    accuracy                           0.74      1331
   macro avg       0.74      0.74      0.74      1331
weighted avg       0.74      0.74      0.74      1331
```



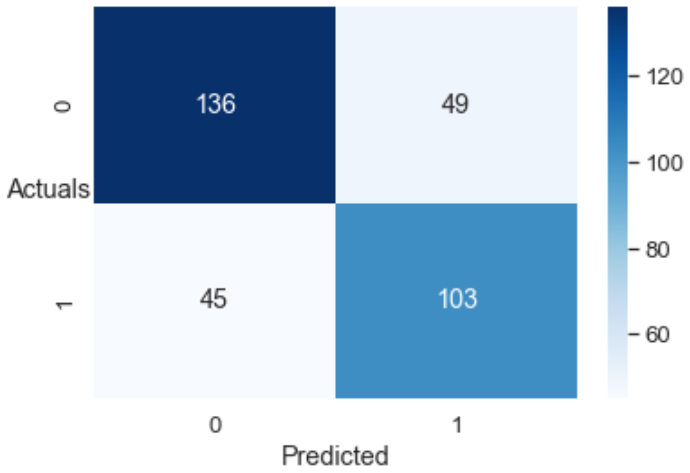
```
The AUC score for Logistic Regression Training dataset is: 0.8296
```



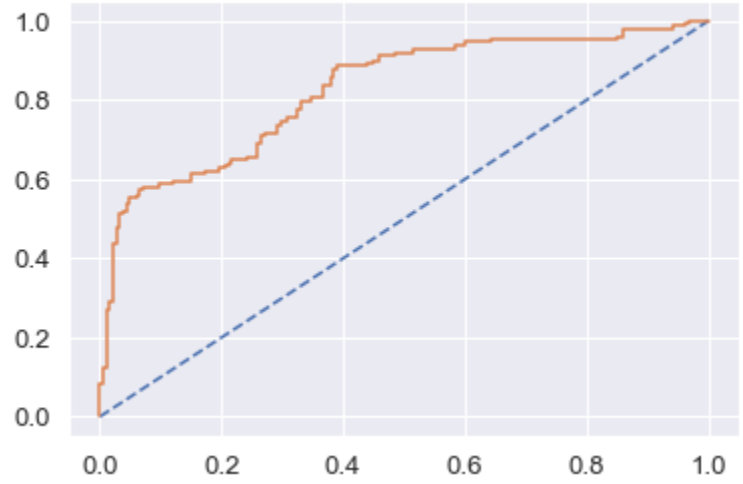
Laptop Logistic Regression Testing Set

```
The model score for Logistic Regression Testing set is 0.7177177177177178
```

```
The Classification Report & Confusion Matrix for Logistic Regression testing set is
              precision    recall  f1-score   support

         0.0       0.75      0.74      0.74       185
         1.0       0.68      0.70      0.69       148

    accuracy                           0.72       333
   macro avg       0.71      0.72      0.71       333
weighted avg       0.72      0.72      0.72       333
```

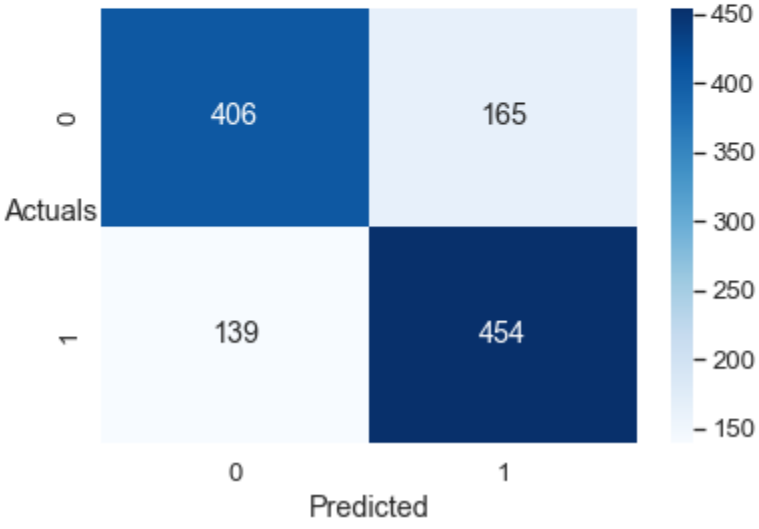


```
The AUC score for Logistic Regression testing set is: 0.8240
```

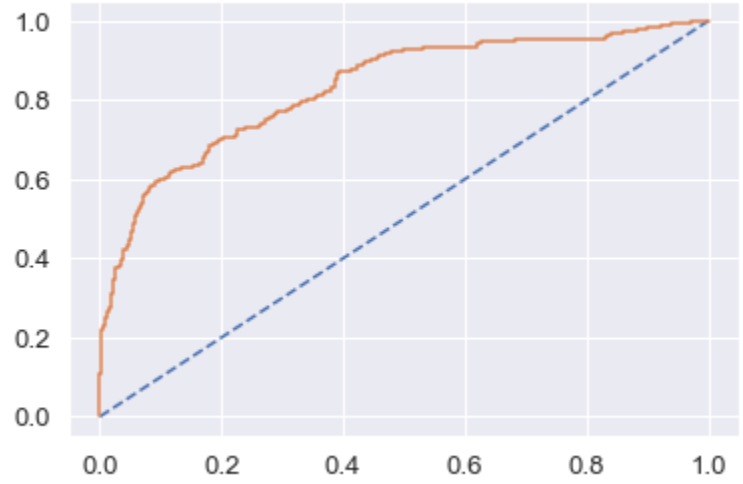


Laptop Linear Discriminant Analysis Training Set

```
The model score for Linear Discriminant Analysis training set is 0.738831615120275

The classification report for Linear Discriminant Analysis training set is
              precision    recall  f1-score   support

         0.0       0.74      0.71      0.73       571
         1.0       0.73      0.77      0.75       593

    accuracy                           0.74      1164
   macro avg       0.74      0.74      0.74      1164
weighted avg       0.74      0.74      0.74      1164
```



```
The AUC score for Linear Discriminant Analysis training set is: 0.831
```

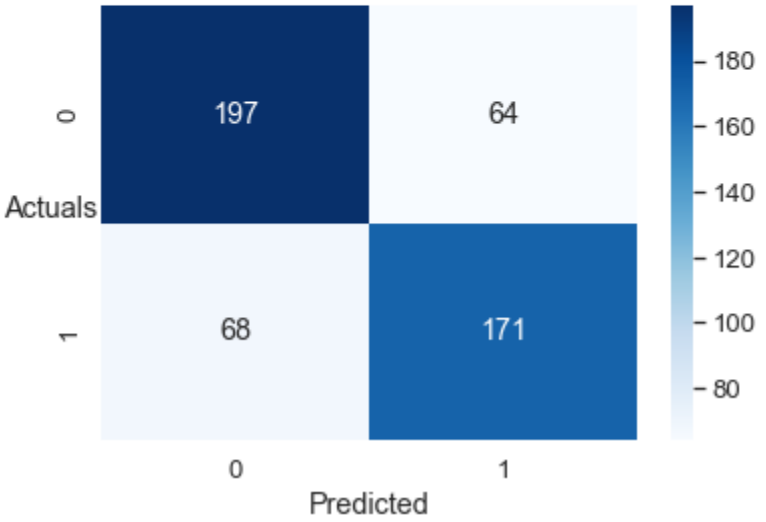


Laptop Linear Discriminant Analysis Testing Set

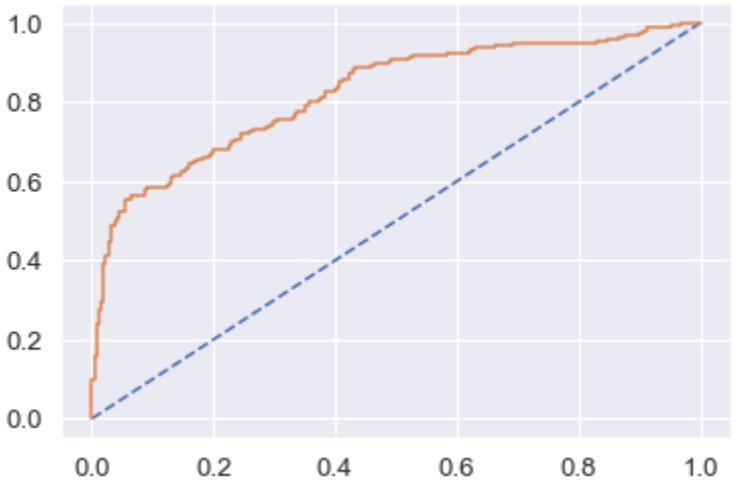
```
The model score for Linear Discriminant Analysis testing set is 0.736
```

The classification report for Linear Discriminant Analysis testing set is

	precision	recall	f1-score	support
0.0	0.74	0.75	0.75	261
1.0	0.73	0.72	0.72	239
accuracy			0.74	500
macro avg	0.74	0.74	0.74	500
weighted avg	0.74	0.74	0.74	500

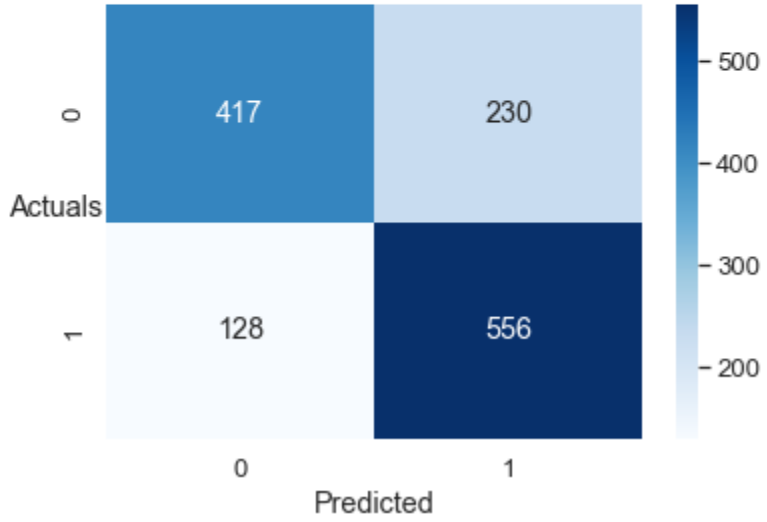


The AUC score for Linear Discriminant Analysis testing set is: 0.822

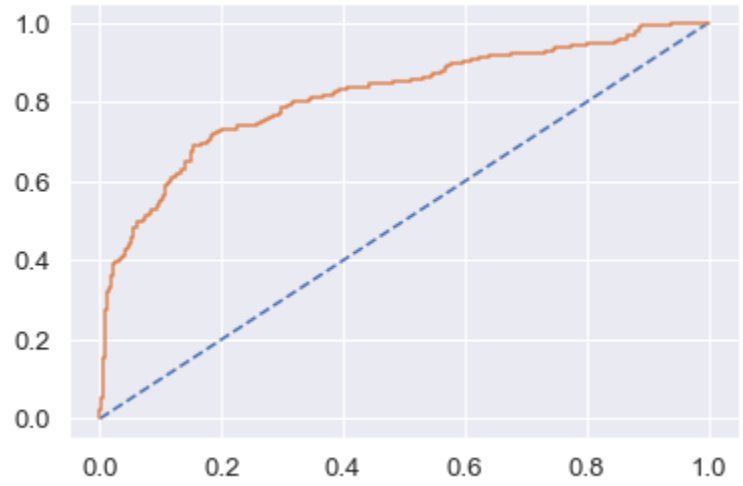


Laptop Naive Bayes Model Training Set

```
The model score for Naive Bayes Model training set is 0.7310293012772352

The classification report for Naive Bayes Model set is
              precision    recall  f1-score   support

         0.0       0.77      0.64      0.70       647
         1.0       0.71      0.81      0.76       684

    accuracy                           0.73      1331
   macro avg       0.74      0.73      0.73      1331
weighted avg       0.74      0.73      0.73      1331
```



```
The AUC score for Naive Bayes training set is: 0.815
```

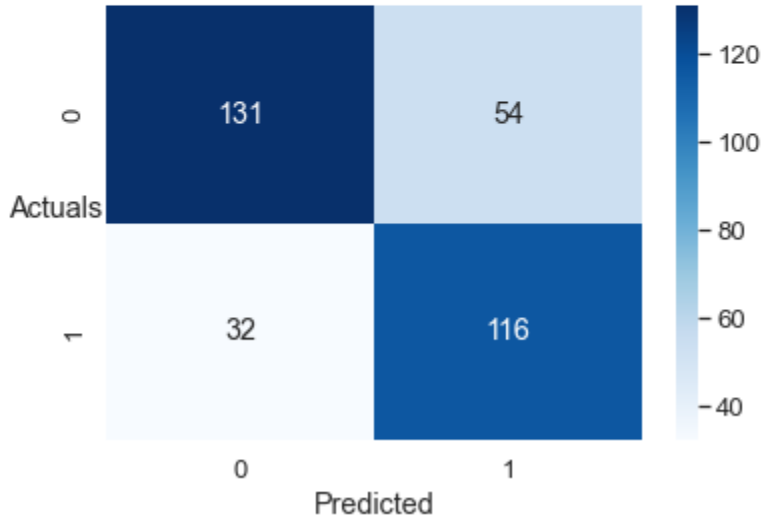


Laptop Naive Bayes Model Testing Set

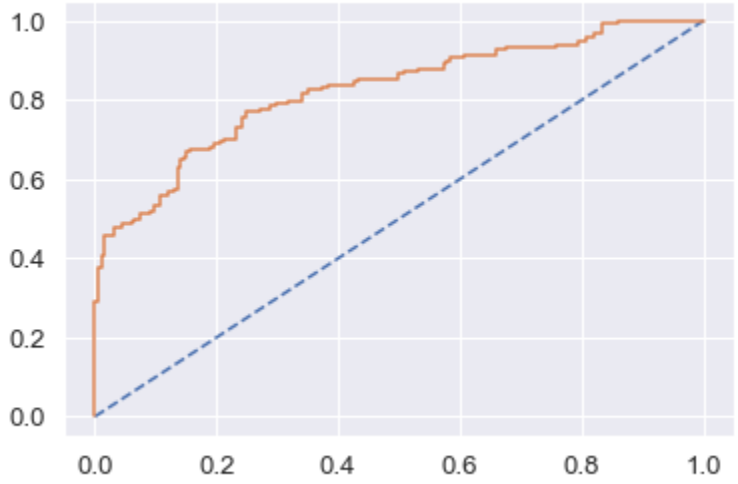
```
The model score for Naive Bayes Model testing set is 0.7417417417417418
```

The classification report for Naive bayes Model testing set is

	precision	recall	f1-score	support
0.0	0.80	0.71	0.75	185
1.0	0.68	0.78	0.73	148
accuracy			0.74	333
macro avg	0.74	0.75	0.74	333
weighted avg	0.75	0.74	0.74	333

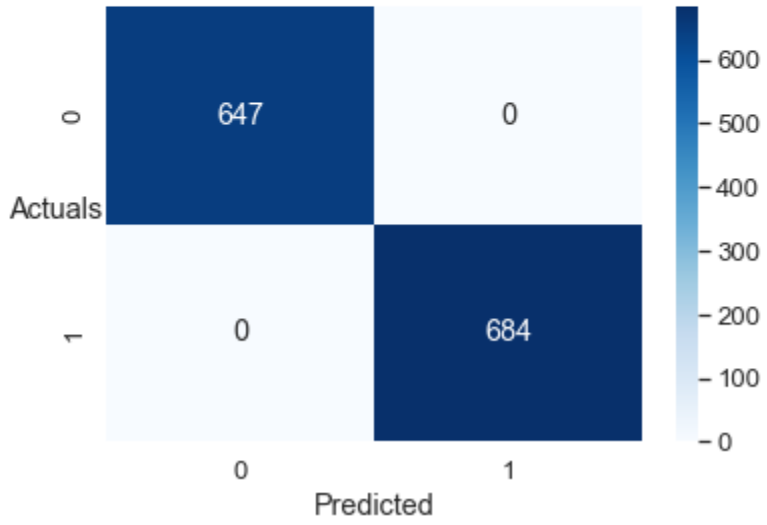


The AUC score for Naive Bayes testing set is: 0.823

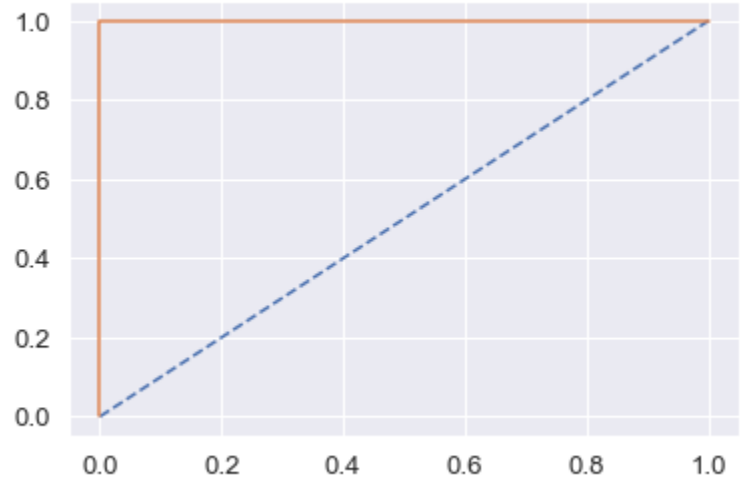


Laptop Decision Tree Classifier Training Set

```
The model score for Decision Tree Classifier training set is 1.0

The classification report for Decision Tree training set is
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00       647
         1.0       1.00      1.00      1.00       684

    accuracy                           1.00      1331
   macro avg       1.00      1.00      1.00      1331
weighted avg       1.00      1.00      1.00      1331
```



```
The AUC score for Decision Tree training set is: 1.000
```

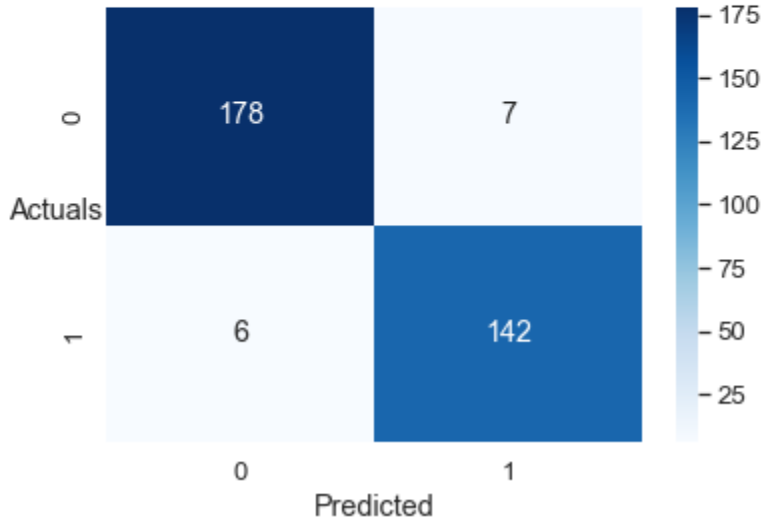


Laptop Decision Tree Classifier Testing Set

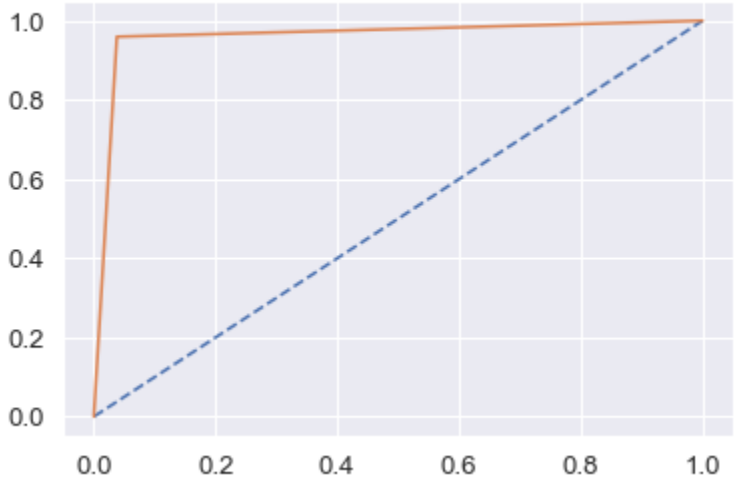
```
The model score for Decision Tree Classifier testing set is 0.960960960960961
```

The classification report for Decision Tree testing set is

	precision	recall	f1-score	support
0.0	0.97	0.96	0.96	185
1.0	0.95	0.96	0.96	148
accuracy			0.96	333
macro avg	0.96	0.96	0.96	333
weighted avg	0.96	0.96	0.96	333

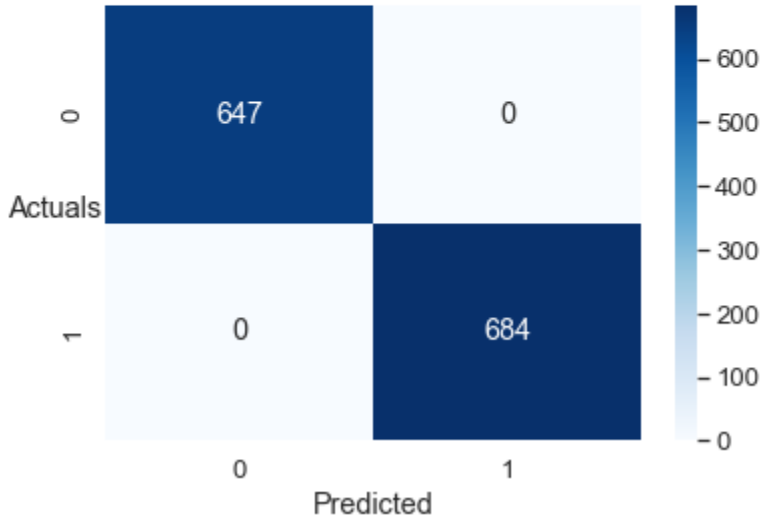


The AUC score for Decision Tree testing set is: 0.961

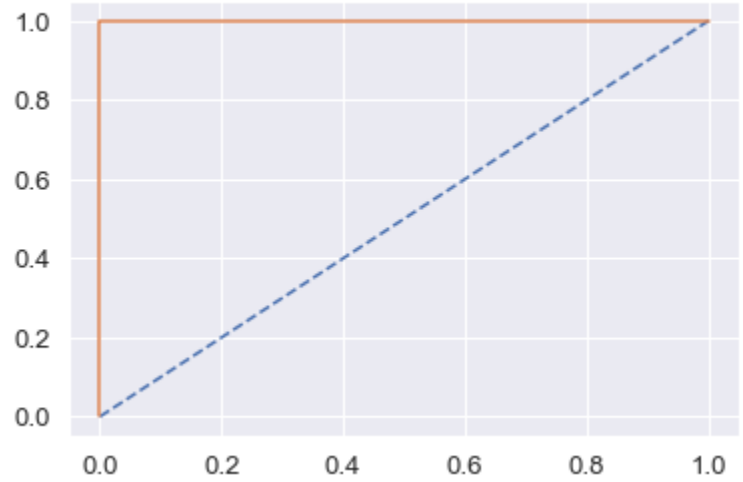


Laptop Random Forest Classifier Training Set

```
The model score for Random Forest Classifier training set is 1.0

The classification report for RFC training set is
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00       647
         1.0       1.00      1.00      1.00       684

    accuracy                           1.00      1331
   macro avg       1.00      1.00      1.00      1331
weighted avg       1.00      1.00      1.00      1331
```

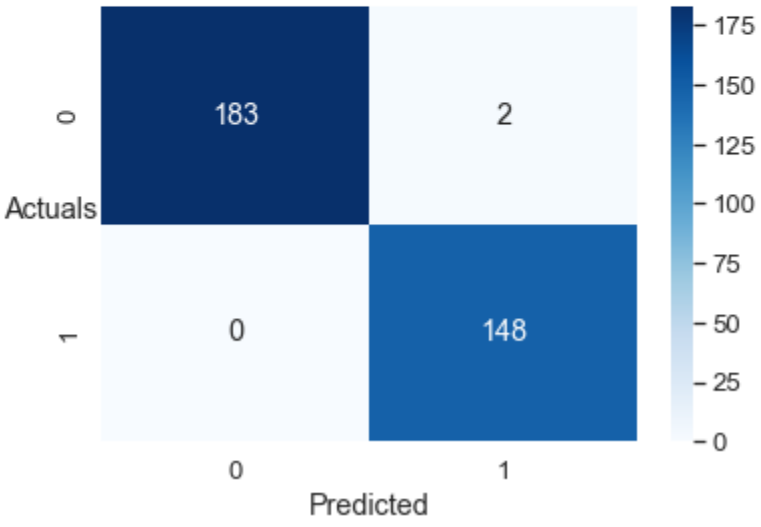


```
The AUC score for RFC training set is: 1.000
```

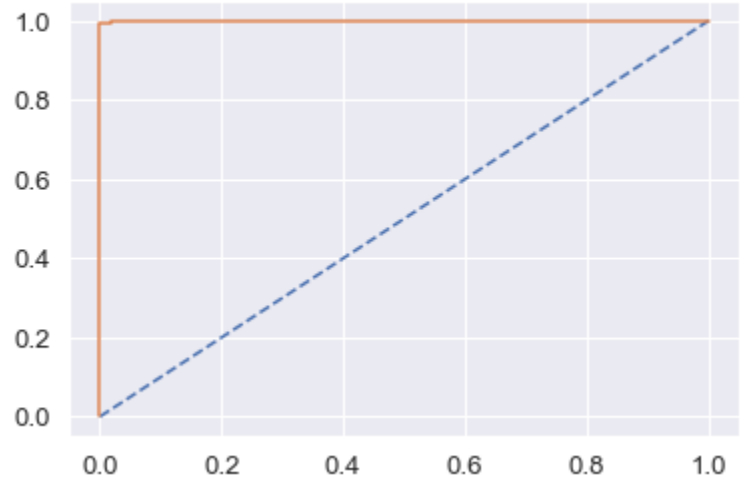


Random Forest Classifier Testing Set

```
The model score for Random Forest Classifier testing set is 0.993993993993994
The classification report for RFC testing set is
              precision    recall  f1-score   support

         0.0       1.00      0.99      0.99       185
         1.0       0.99      1.00      0.99       148

    accuracy                           0.99       333
   macro avg       0.99      0.99      0.99       333
weighted avg       0.99      0.99      0.99       333
```



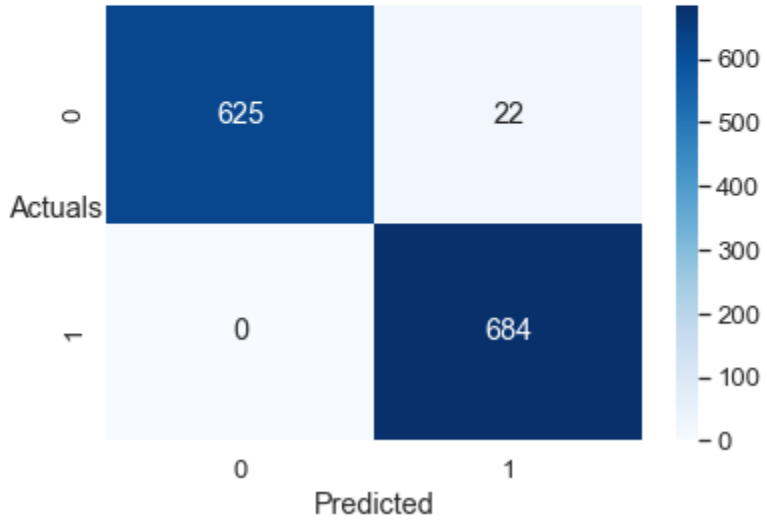
```
The AUC score for RFC testing set is: 1.000
```



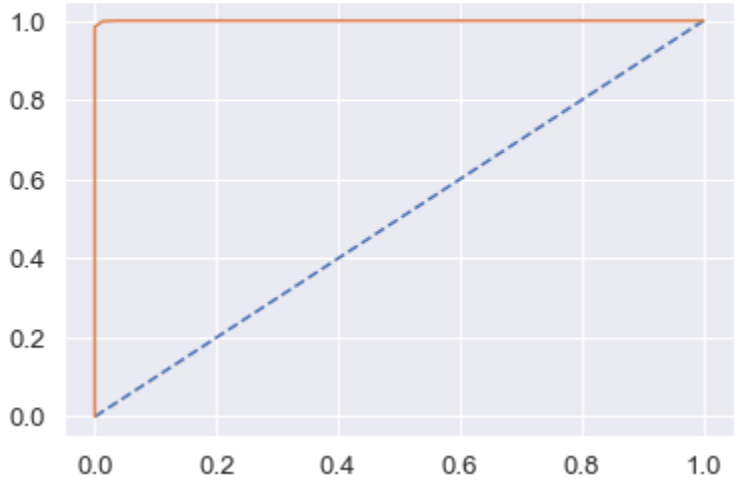
Laptop K- Nearest Neighbour Training Set

```
The model score for KNN training set is 0.9834710743801653
```

```
The classification report for KNN set is
              precision    recall  f1-score   support

         0.0       1.00      0.97      0.98       647
         1.0       0.97      1.00      0.98       684

    accuracy                           0.98      1331
   macro avg       0.98      0.98      0.98      1331
weighted avg       0.98      0.98      0.98      1331
```




```
The AUC score for KNN training set is: 1.000
```

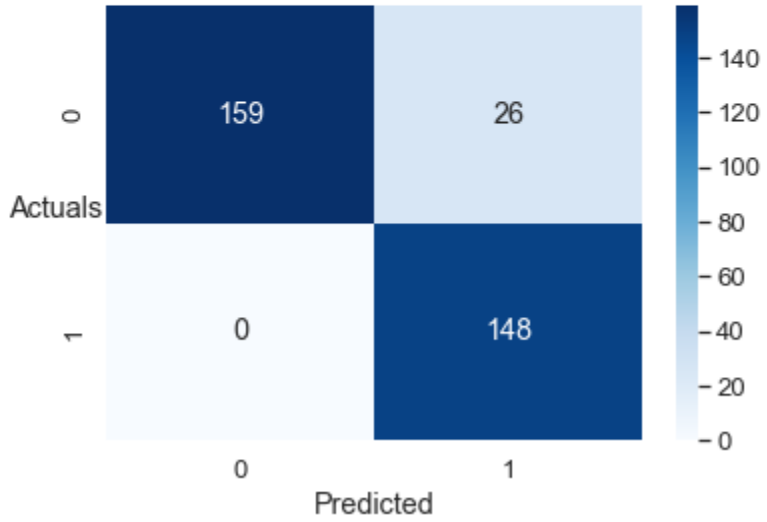


Laptop K- Nearest Neighbour Testing Set

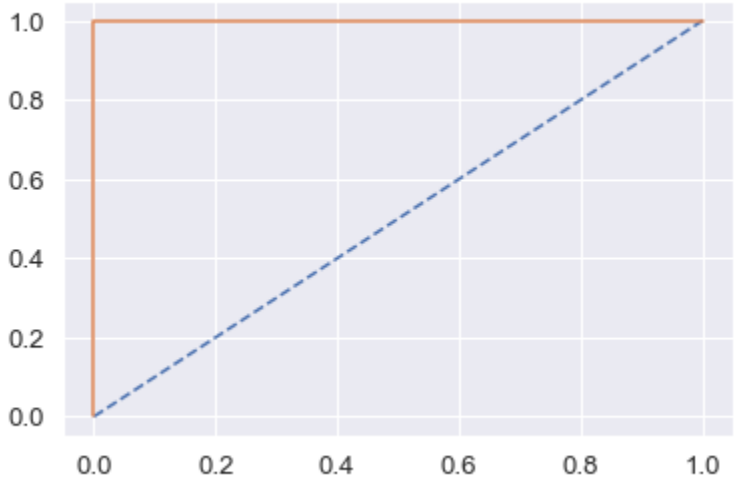
```
The model score for KNN testing set is 0.9219219219219219
```

The classification report for KNN testing set is

	precision	recall	f1-score	support
0.0	1.00	0.86	0.92	185
1.0	0.85	1.00	0.92	148
accuracy			0.92	333
macro avg	0.93	0.93	0.92	333
weighted avg	0.93	0.92	0.92	333



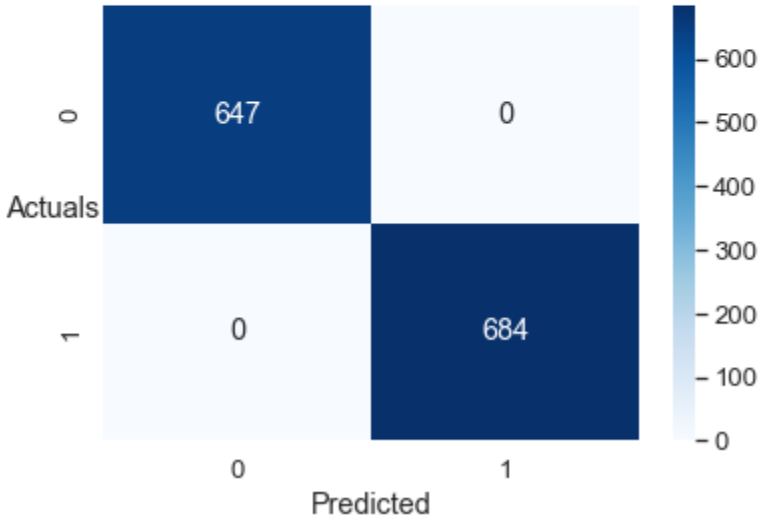
```
The AUC score for KNN testing set is: 1.000
```



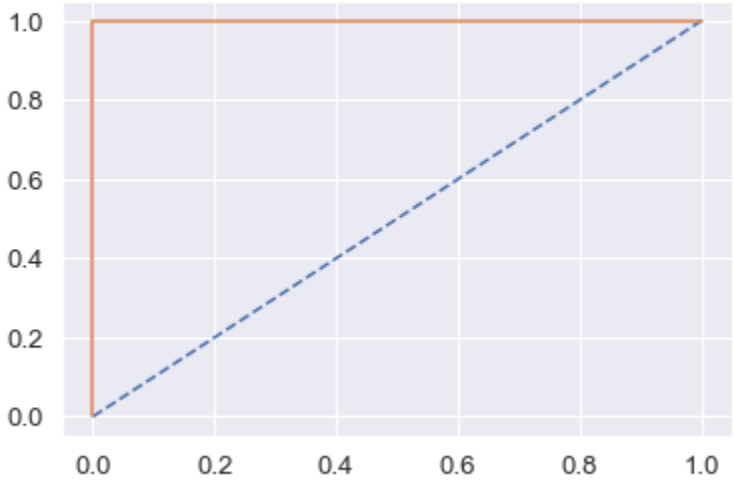
2. Model Tuning and business implications

Laptop Bagging training set

```
The model score for Bagging training set is 1.0

The classification report for Bagging training set is
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00       647
         1.0       1.00      1.00      1.00       684

    accuracy                           1.00      1331
   macro avg       1.00      1.00      1.00      1331
weighted avg       1.00      1.00      1.00      1331
```



```
The AUC score for Bagging training set is: 1.000
```

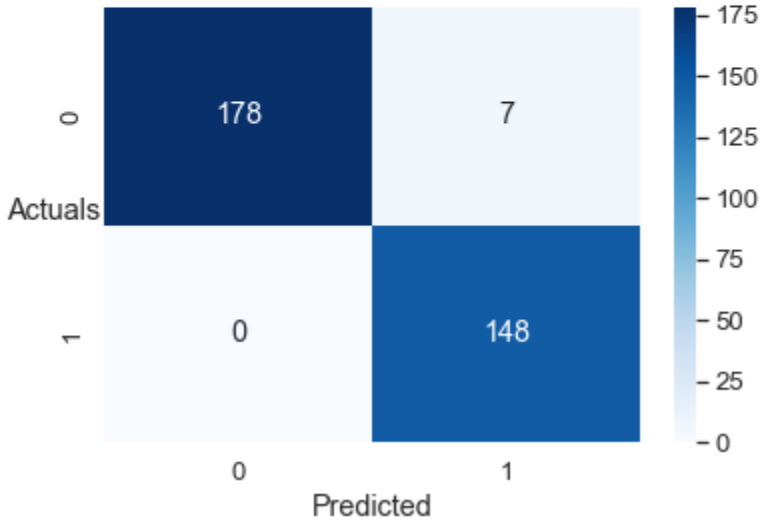


Laptop Bagging testing set

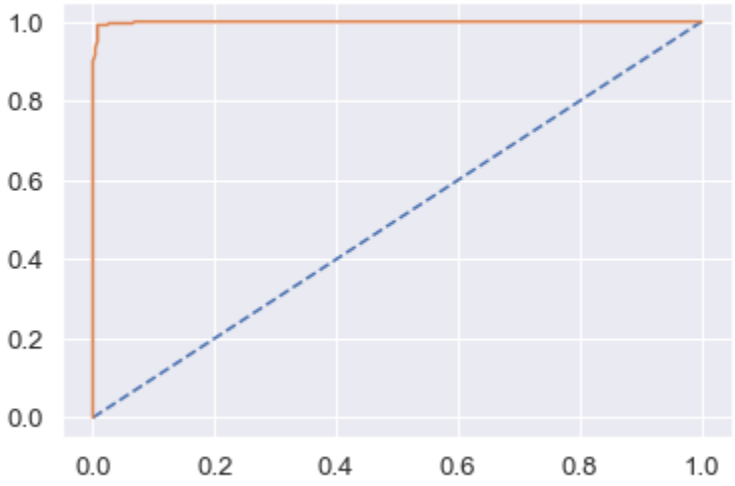
```
The model score for Bagging testing set is 0.978978978978979
```

The classification report for Bagging testing set is

	precision	recall	f1-score	support
0.0	1.00	0.96	0.98	185
1.0	0.95	1.00	0.98	148
accuracy			0.98	333
macro avg	0.98	0.98	0.98	333
weighted avg	0.98	0.98	0.98	333

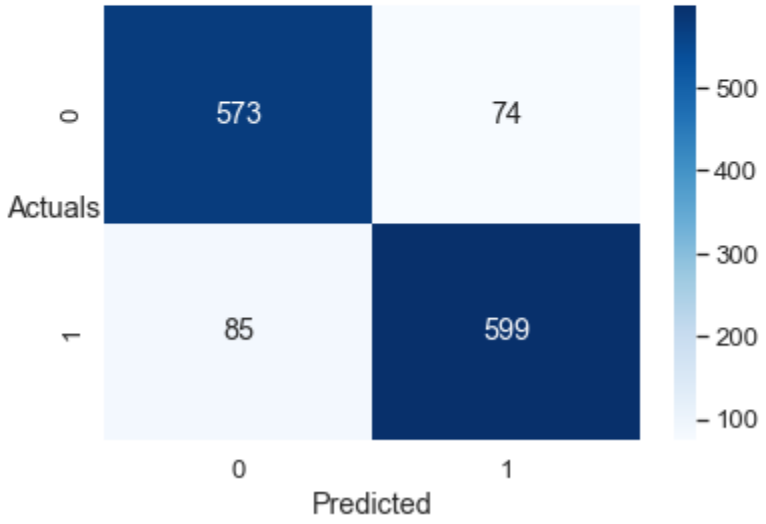


The AUC score for Bagging testing set is: 0.999

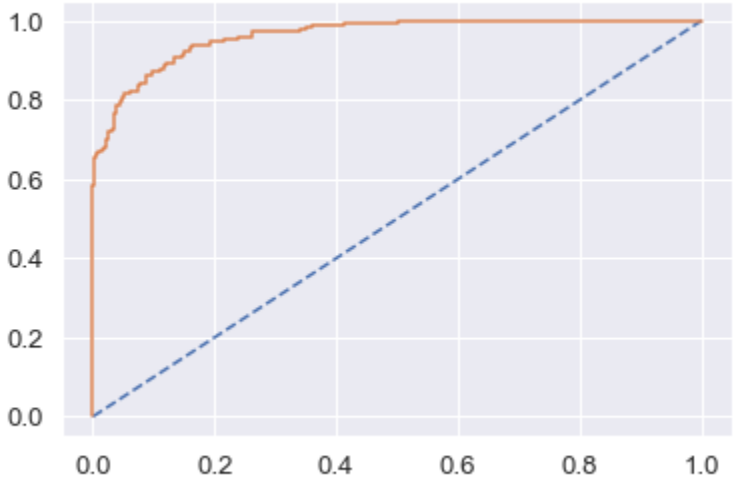


Laptop AdaBoosting on Training Set

```
The model score for AdaBoosting training set is 0.8805409466566492

The classification report for Adaboosting training set is
              precision    recall  f1-score   support

         0.0       0.87      0.89      0.88       647
         1.0       0.89      0.88      0.88       684

    accuracy                           0.88      1331
   macro avg       0.88      0.88      0.88      1331
weighted avg       0.88      0.88      0.88      1331
```



```
The AUC score for AdaBoosting training set is: 0.962
```

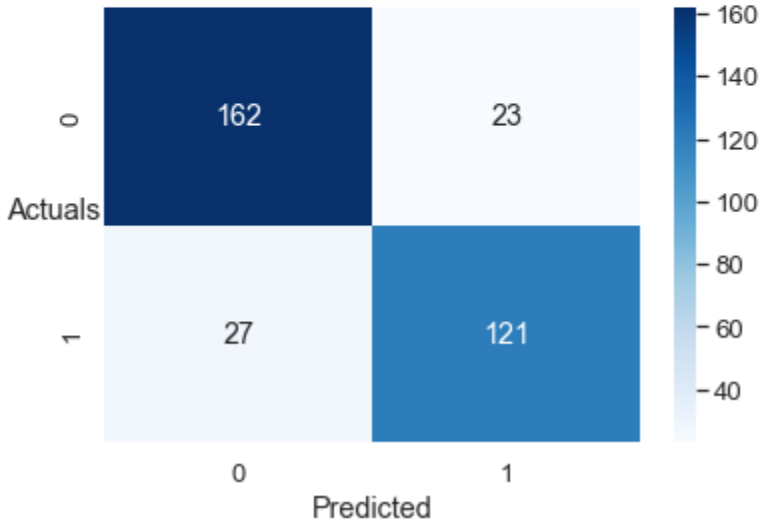


Laptop AdaBoosting on Testing Set

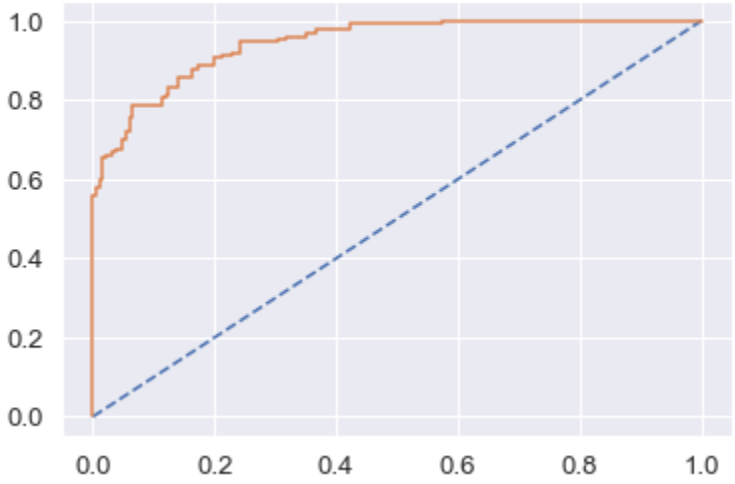
```
The model score for AdaBoosting testing set is 0.8498498498498499
```

The classification report for Adaboosting testing set is

	precision	recall	f1-score	support
0.0	0.86	0.88	0.87	185
1.0	0.84	0.82	0.83	148
accuracy			0.85	333
macro avg	0.85	0.85	0.85	333
weighted avg	0.85	0.85	0.85	333

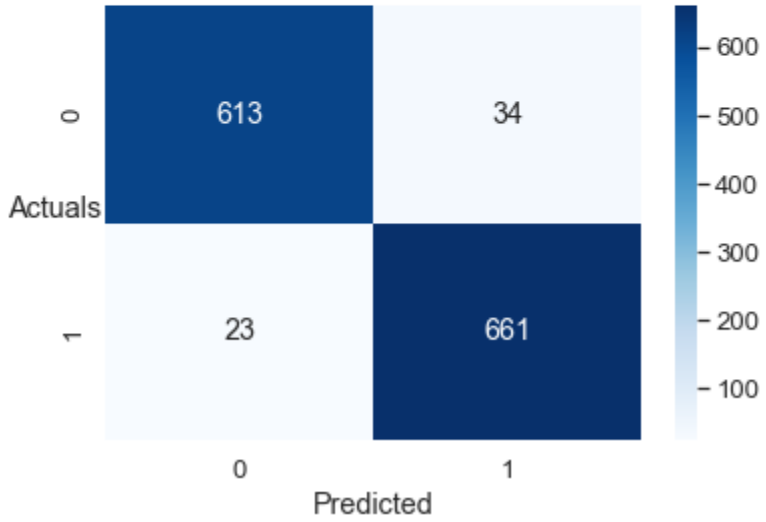


The AUC score for AdaBoosting testing set is: 0.943

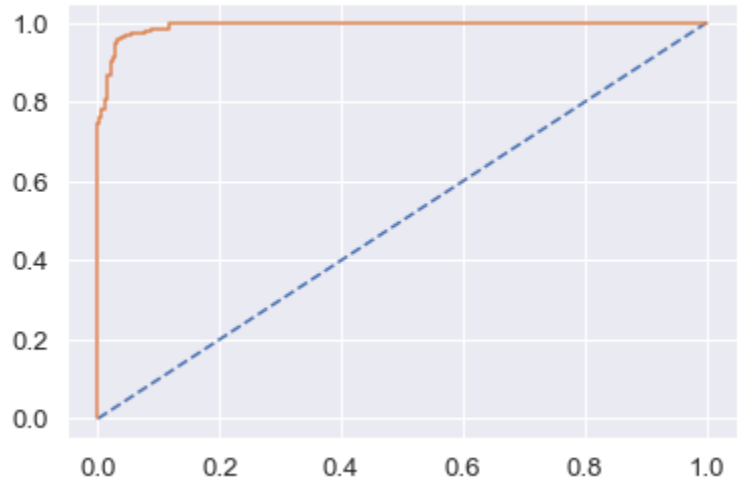


Laptop Gradient Boosting Training Set

```
The model score for GradientBoosting training set is 0.9571750563486101

The classification report for Gradientboosting training set is
              precision    recall  f1-score   support

         0.0       0.96      0.95      0.96       647
         1.0       0.95      0.97      0.96       684

    accuracy                           0.96      1331
   macro avg       0.96      0.96      0.96      1331
weighted avg       0.96      0.96      0.96      1331
```




```
The AUC score for GradientBoosting training set is: 0.992
```

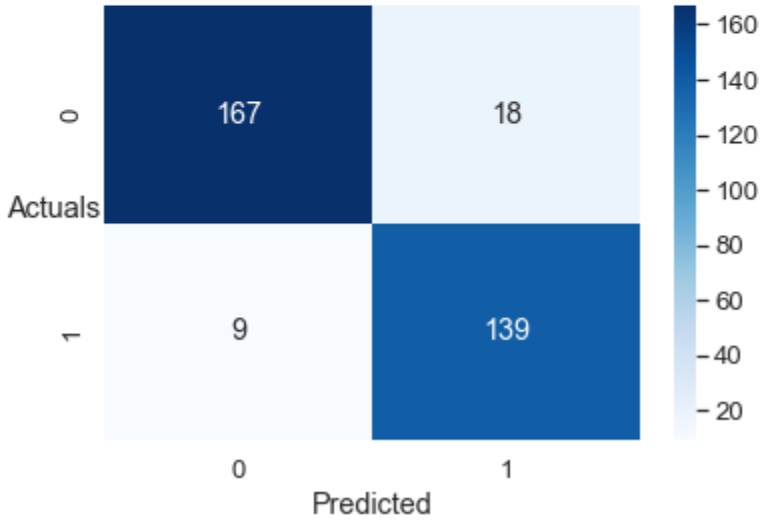


Laptop Gradient Boosting Testing Set

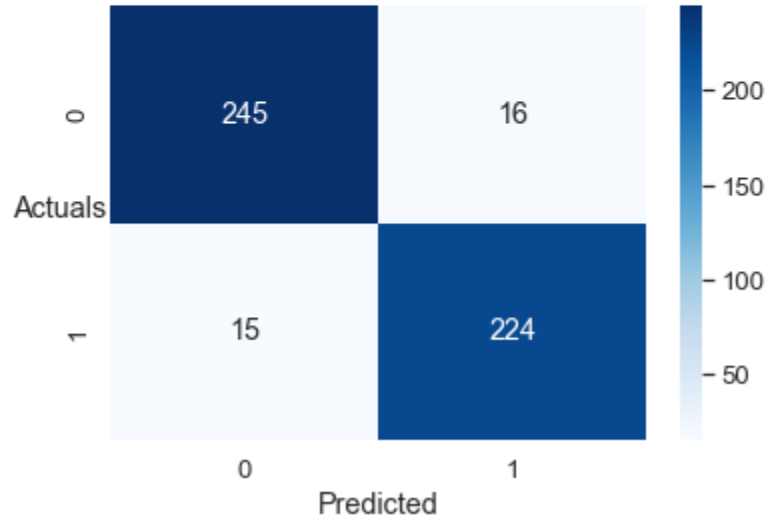
```
The model score for GradientBoosting testing set is 0.918918918918919
```

The classification report for Gradientboosting testing set is

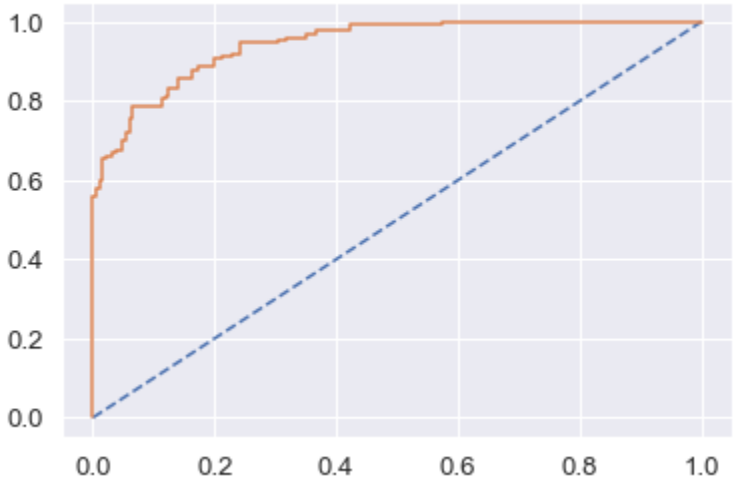
	precision	recall	f1-score	support
0.0	0.95	0.90	0.93	185
1.0	0.89	0.94	0.91	148
accuracy			0.92	333
macro avg	0.92	0.92	0.92	333
weighted avg	0.92	0.92	0.92	333



```
The classification report for Gradientboosting testing set is
              precision    recall  f1-score   support

         0.0       0.94      0.94      0.94       261
         1.0       0.93      0.94      0.94       239

    accuracy                           0.94       500
   macro avg       0.94      0.94      0.94       500
weighted avg       0.94      0.94      0.94       500
```




```
The AUC score for GradientBoosting testing set is: 0.943
```



Comparing all the models

	LR Train	LR Test	LDA Train	LDA Test	NB Train	NB Test	CART Train	CART Test	RFC Train	RFC Test	Bagging Train	Bagging Test	Ada Boosting Train	Ada Boosting Test	KNN Train	KNN Test	Gradient Boosting Train	Gradient Boosting Test
Precision	0.739	0.678	0.734	0.667	0.707	0.682	1.0	0.953	1.0	0.987	1.0	0.955	0.890	0.840	0.969	0.851	0.951	0.885
Recall	0.756	0.696	0.770	0.703	0.813	0.784	1.0	0.959	1.0	1.000	1.0	1.000	0.876	0.818	1.000	1.000	0.966	0.939
F1 Score	0.747	0.687	0.752	0.684	0.756	0.730	1.0	0.956	1.0	0.993	1.0	0.977	0.883	0.829	0.984	0.919	0.959	0.911
Accuracy	0.737	0.718	0.739	0.712	0.731	0.742	1.0	0.961	1.0	0.994	1.0	0.979	0.881	0.850	0.983	0.922	0.957	0.919
AUC Score	0.830	0.824	0.829	0.822	0.815	0.823	1.0	0.961	1.0	1.000	1.0	0.999	0.962	0.943	1.000	1.000	0.992	0.943

Inferences for Laptop devices based on Model Building

i. Logistic Regression

The Logistic Regression has come up with poor accuracy of 73.7% on the train set and only 71.8% accuracy on test set. The precision also is not up to the mark i.e. 73.9% for train and 67.8% for the test datasets.

ii. Linear Discriminant Analysis

The LDA has also not performed well. It has given an accuracy score of 73.9% on Train set and 73.9% on the Test set. The precision also isn't up to the mark. Precision is only 73.4% for train set and 66.7 for the test set.

iii. Naïve Bayes Model

The Naïve Bayes model has also not shown very poor performance. The accuracy score is 73.1% and 74.2% on train as well as test split datasets. The precision is poor for this model. It is 70.7% & 68.2% for train as well as test sets.

iv. Decision Tree Classifier

The Decision Tree Classifier (CART) model has performed reasonably well. The accuracy score for train and test dataset is 100% and 96.1%. The precision is also very good. It is 100% and 95.3% for train and test set respectively.

v. Random Forest Classifier

The performance of the Random Forest Classifier is the best amongst all the models. The accuracy is 100% for the train and

99.4% for the test data sets. The precision is also 100% for training dataset and 98.7% testing dataset.

vi. K- Nearest Neighbour

This model has performed fairly but not up to the mark. The accuracy is 98.3% for train and 92.2% for the test dataset.

Model Tuning (Bagging & Boosting)

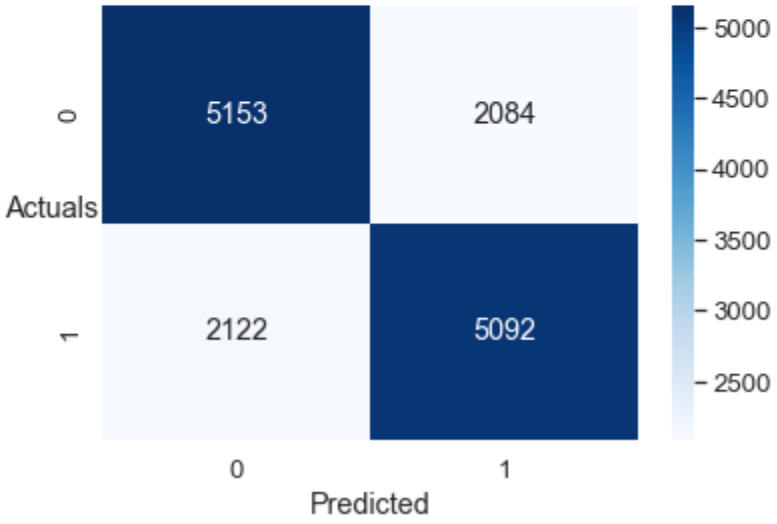
After applying the model tuning technique bagging to the model the performance received was good. Accuracy for train dataset was 100% and 97.9% for the test. In the case of boosting the Gradient boosting gave accuracy of 95.7% on train and 91.9% on the test datasets.

The final recommendation to business shall be move ahead with Random Forest Classifier, where on the test dataset the False Negatives were 0 and False Positives were only 2.

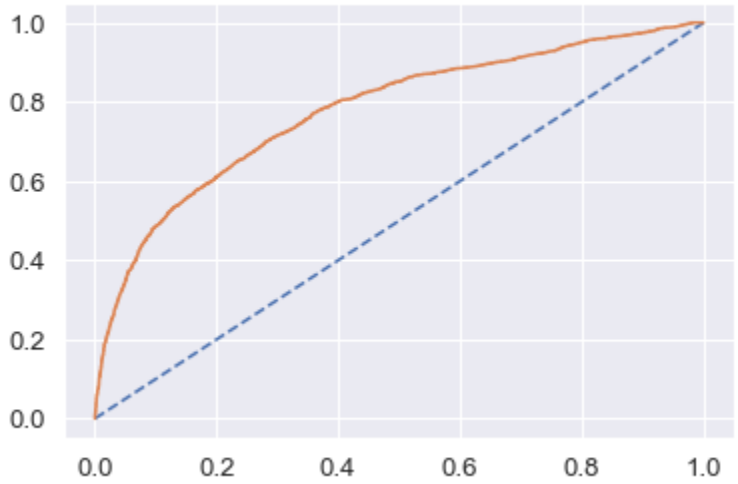
MOBILE-

Mobile Logistic Regression Training Set

```
The model score for Logistic Regression training set is 0.7089474776832053

The classification report for Logistic Regression training set is
              precision    recall  f1-score   support

         0.0       0.71      0.71      0.71      7237
         1.0       0.71      0.71      0.71      7214

    accuracy                           0.71     14451
   macro avg       0.71      0.71      0.71     14451
weighted avg       0.71      0.71      0.71     14451
```



```
The AUC score for Logistic Regression training set is: 0.777
```

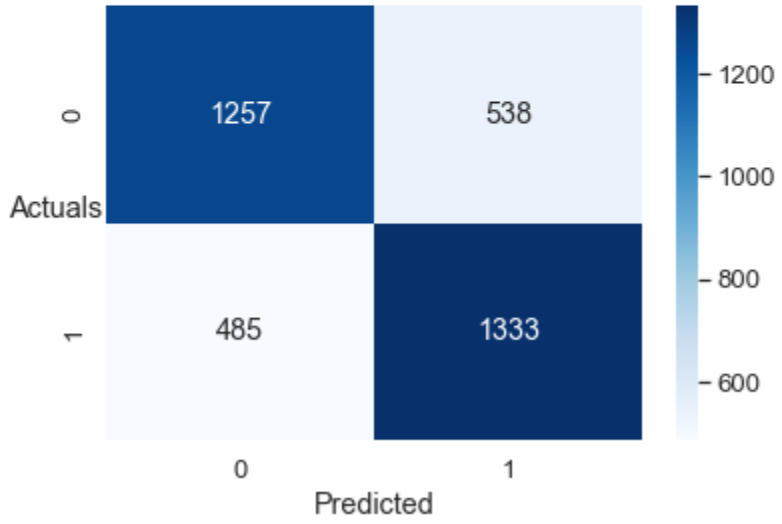


Mobile Logistic Regression Testing Set

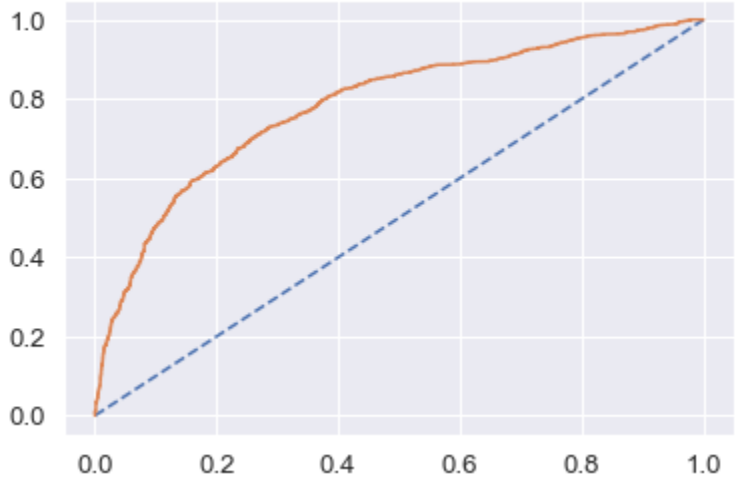
```
The model score for Logistic Regression testing set is 0.7168557985053972
```

The classification report for Logistic Regression testing set is

	precision	recall	f1-score	support
0.0	0.72	0.70	0.71	1795
1.0	0.71	0.73	0.72	1818
accuracy			0.72	3613
macro avg	0.72	0.72	0.72	3613
weighted avg	0.72	0.72	0.72	3613

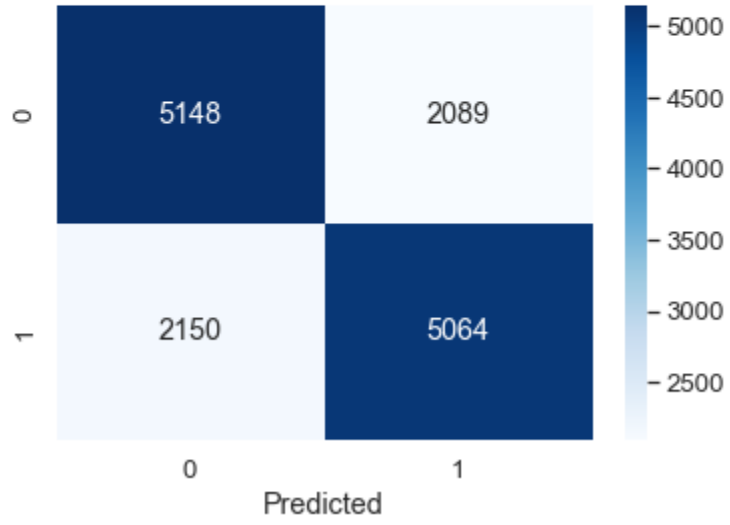


The AUC score for Logistic Regression testing set is: 0.783

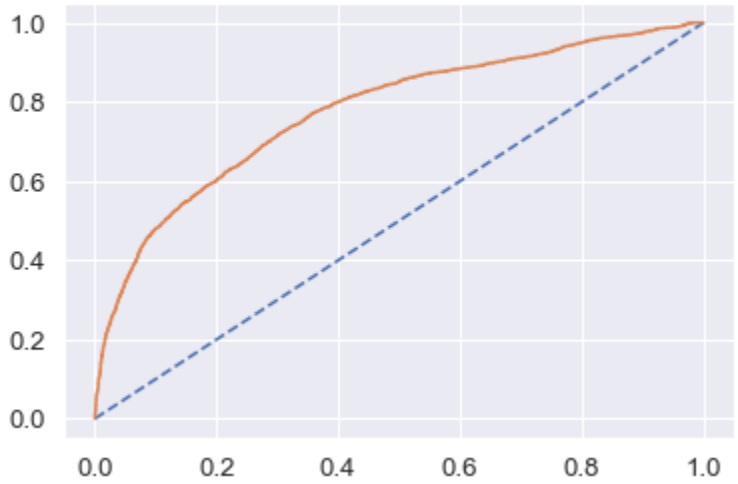


Mobile Linear Discriminant Analysis Training Set

```
The model score for Linear Discriminant Analysis training set is 0.7066638986921321


The classification report for Linear Discriminant Analysis training set is
              precision    recall  f1-score   support

         0.0       0.71      0.71      0.71      7237
         1.0       0.71      0.70      0.70      7214

    accuracy                           0.71     14451
   macro avg       0.71      0.71      0.71     14451
weighted avg       0.71      0.71      0.71     14451
```



```
The AUC score for Linear Discriminant Analysis training set is: 0.776
```

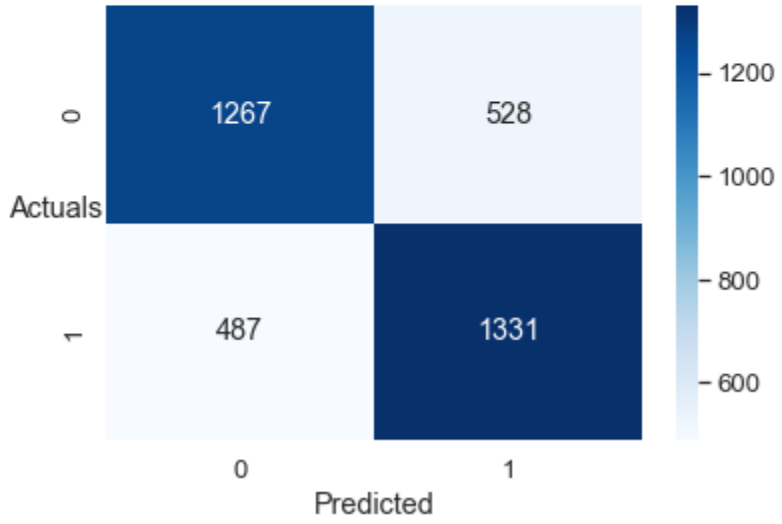


Mobile Linear Discriminant Analysis Testing Set

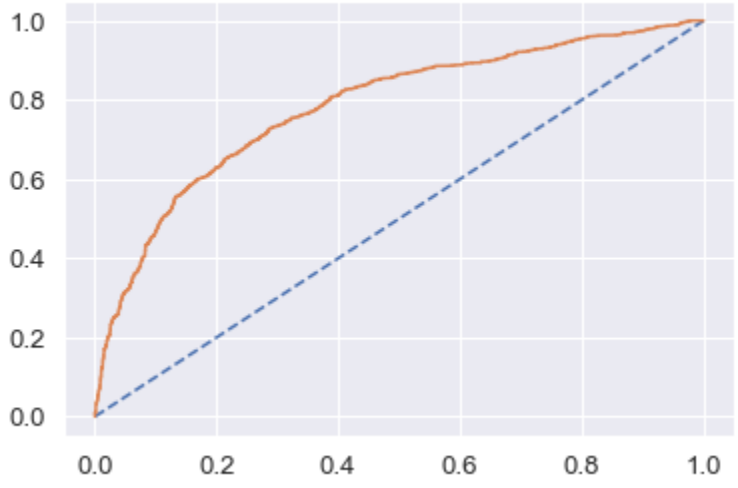
```
The model score for Linear Discriminant Analysis testing set is 0.719070024910047
```

The classification report for Linear Discriminant Analysis testing set is

	precision	recall	f1-score	support
0.0	0.72	0.71	0.71	1795
1.0	0.72	0.73	0.72	1818
accuracy			0.72	3613
macro avg	0.72	0.72	0.72	3613
weighted avg	0.72	0.72	0.72	3613

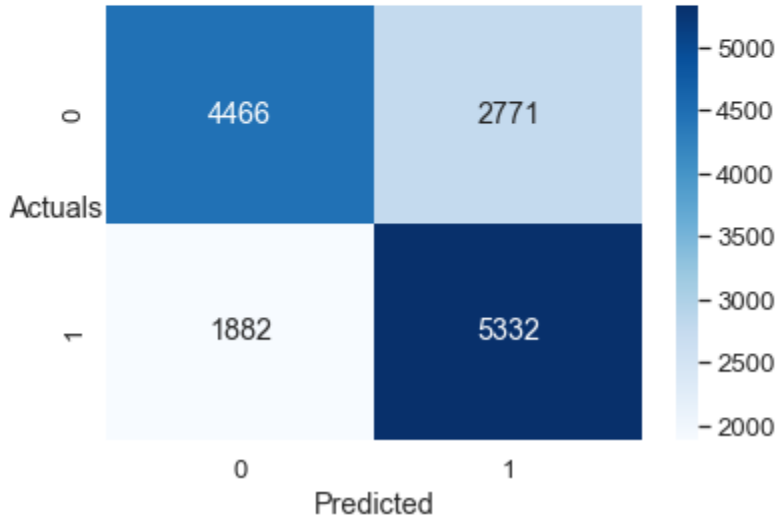


The AUC score for Linear Discriminant Analysis testing set is: 0.782

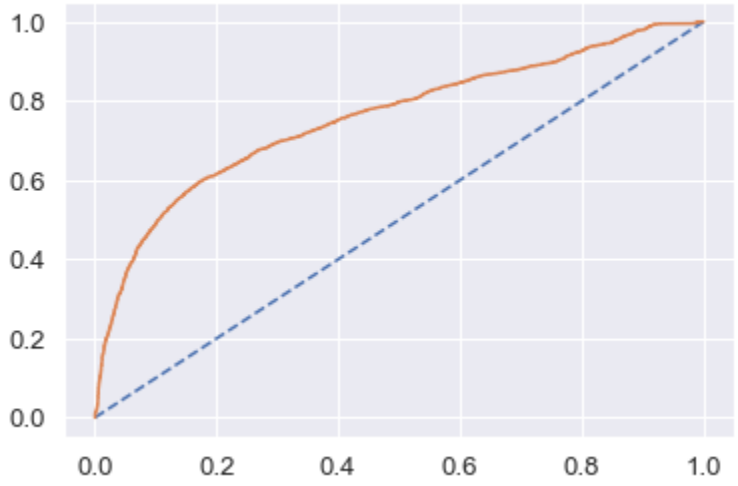


Mobile Naïve Baiyes Training Set

```
The model score for Naive Bayes Model training set is 0.6780153622586672

The classification report for Naive Bayes Model set is
              precision    recall  f1-score   support

         0.0       0.70      0.62      0.66      7237
         1.0       0.66      0.74      0.70      7214

    accuracy                           0.68     14451
   macro avg       0.68      0.68      0.68     14451
weighted avg       0.68      0.68      0.68     14451
```

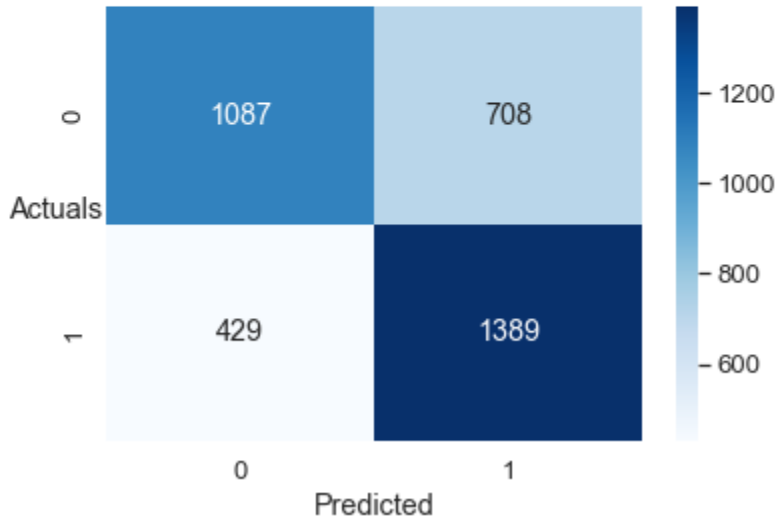



```
The AUC score for Naive Bayes training set is: 0.757
```

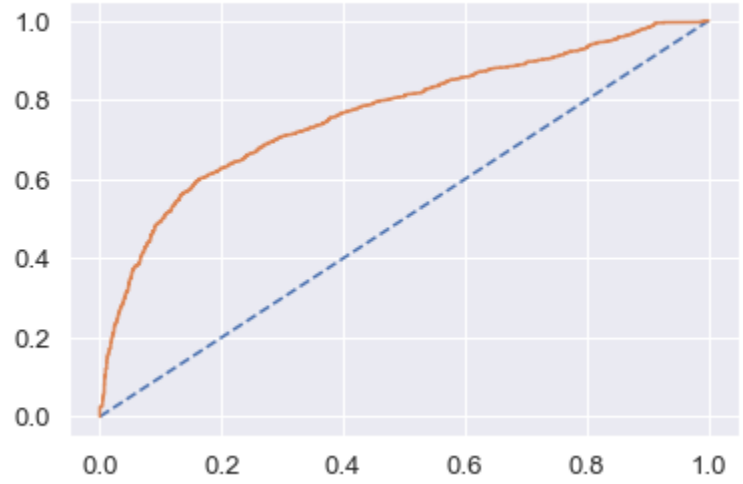


Mobile Naïve Baiyes Testing Set

```
The model score for Naive Bayes Model testing set is 0.6853030722391364

The classification report for Naive bayes Model testing set is
              precision    recall  f1-score   support

         0.0       0.72      0.61      0.66      1795
         1.0       0.66      0.76      0.71      1818

    accuracy                           0.69      3613
   macro avg       0.69      0.68      0.68      3613
weighted avg       0.69      0.69      0.68      3613
```



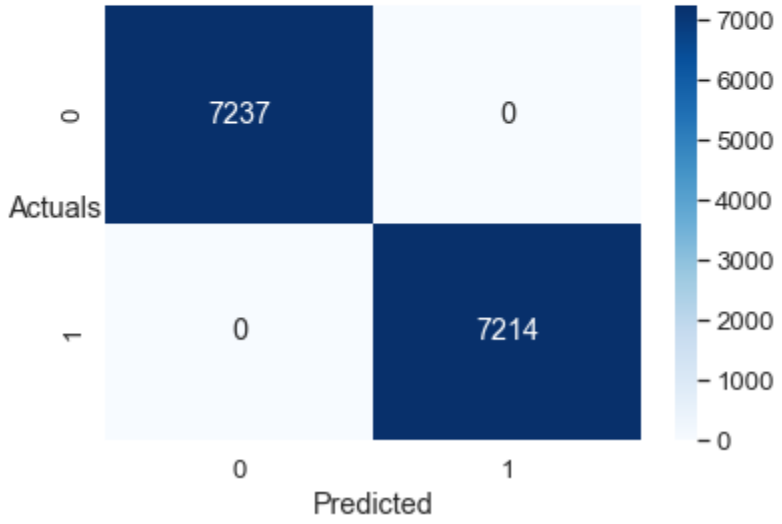
```
The AUC score for Naive Bayes testing set is: 0.765
```



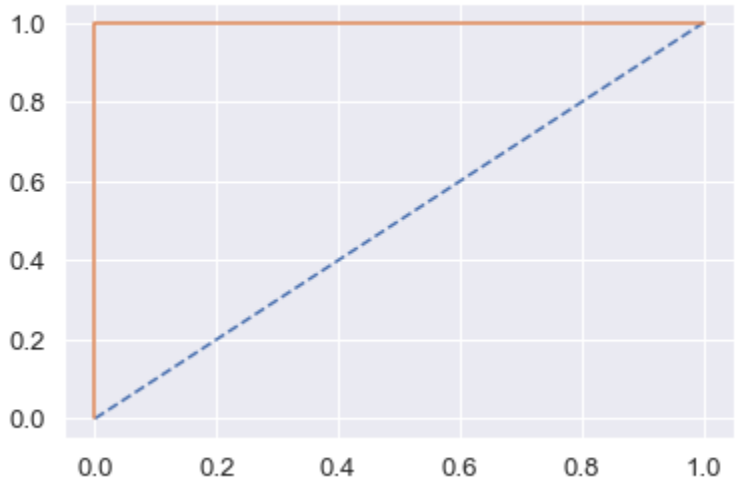
Mobile Decision Tree Classifier Training Set

```
The model score for Decision Tree Classifier training set is 1.0
```

```
The classification report for Decision Tree training set is
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00      7237
         1.0       1.00      1.00      1.00      7214

    accuracy                           1.00     14451
   macro avg       1.00      1.00      1.00     14451
weighted avg       1.00      1.00      1.00     14451
```

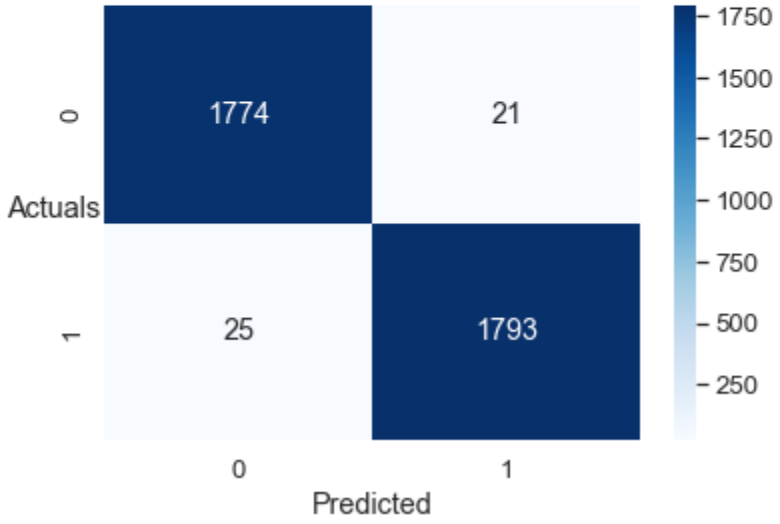



```
The AUC score for Decision Tree training set is: 1.000
```

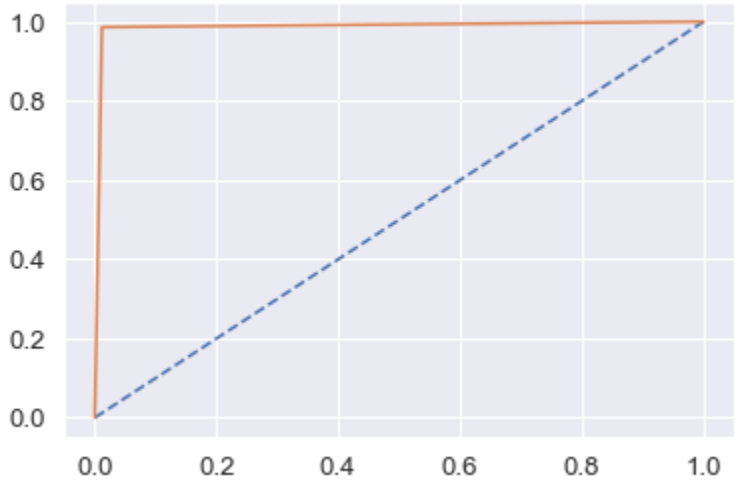


Mobile Decision Tree Classifier Testing Set

```
The model score for Decision Tree Classifier testing set is 0.9872681981732632
The classification report for Decision Tree testing set is
              precision    recall  f1-score   support

         0.0       0.99      0.99      0.99      1795
         1.0       0.99      0.99      0.99      1818

    accuracy                           0.99      3613
   macro avg       0.99      0.99      0.99      3613
weighted avg       0.99      0.99      0.99      3613
```

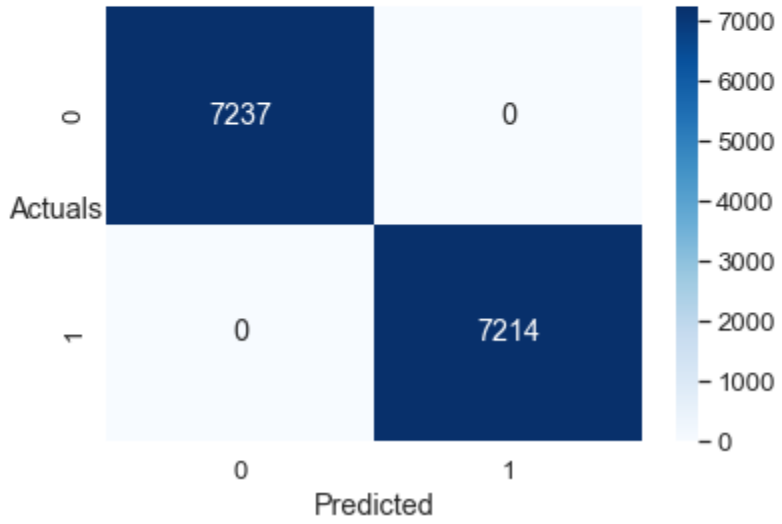



```
The AUC score for Decision Tree testing set is: 0.987
```

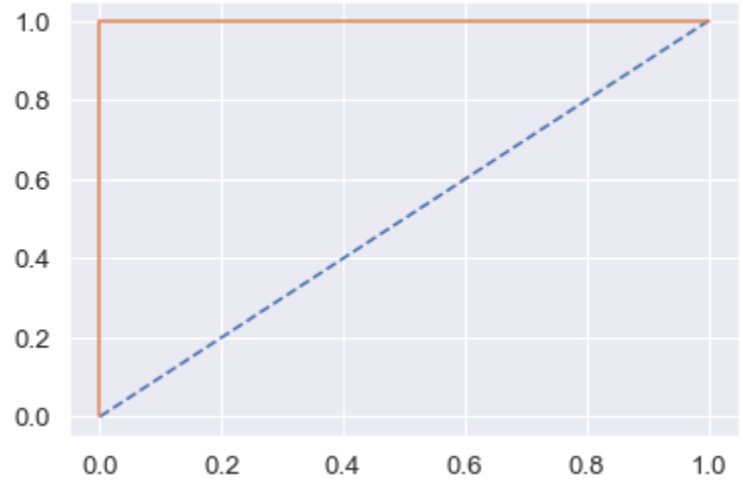


Mobile Random Forest Classifier Training Set

```
The model score for Random Forest Classifier training set is 1.0

The classification report for RFC training set is
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00      7237
         1.0       1.00      1.00      1.00      7214

    accuracy                           1.00     14451
   macro avg       1.00      1.00      1.00     14451
weighted avg       1.00      1.00      1.00     14451
```

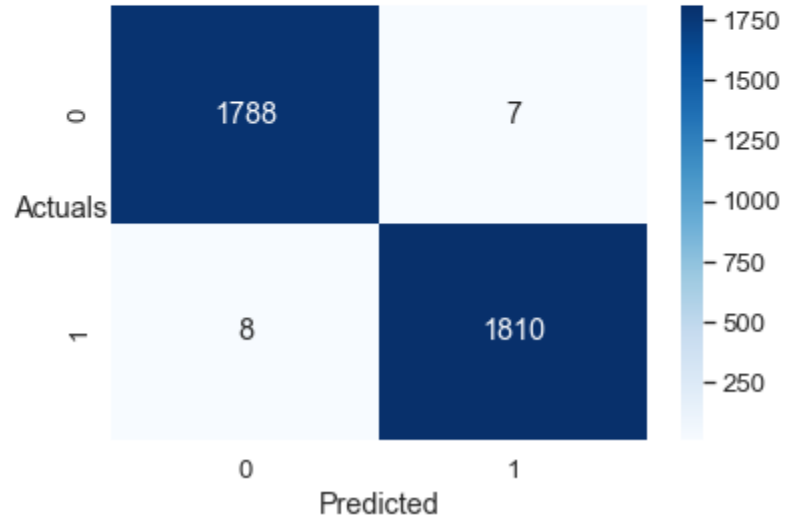


```
The AUC score for RFC training set is: 1.000
```

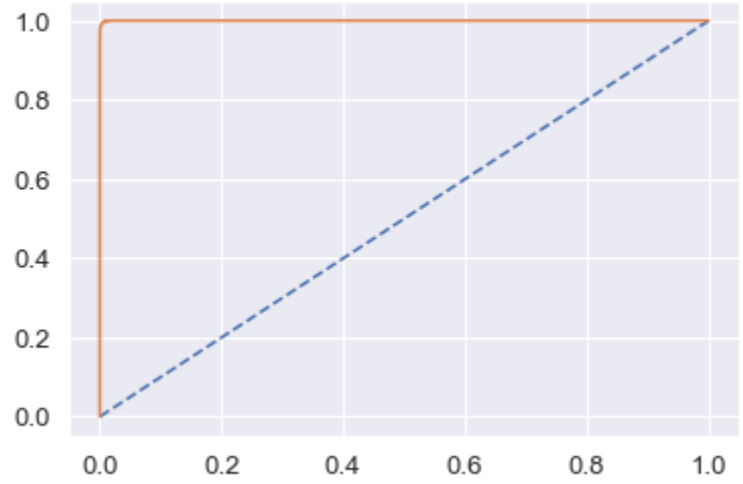


Mobile Random Forest Classifier Testing Set

```
The model score for Random Forest Classifier testing set is 0.9942804428044281

The classification report for RFC testing set is
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00      1795
         1.0       1.00      1.00      1.00      1818

    accuracy                           1.00      3613
   macro avg       1.00      1.00      1.00      3613
weighted avg       1.00      1.00      1.00      3613
```



```
The AUC score for RFC testing set is: 1.000
```



Mobile K Nearest Neighbour Training

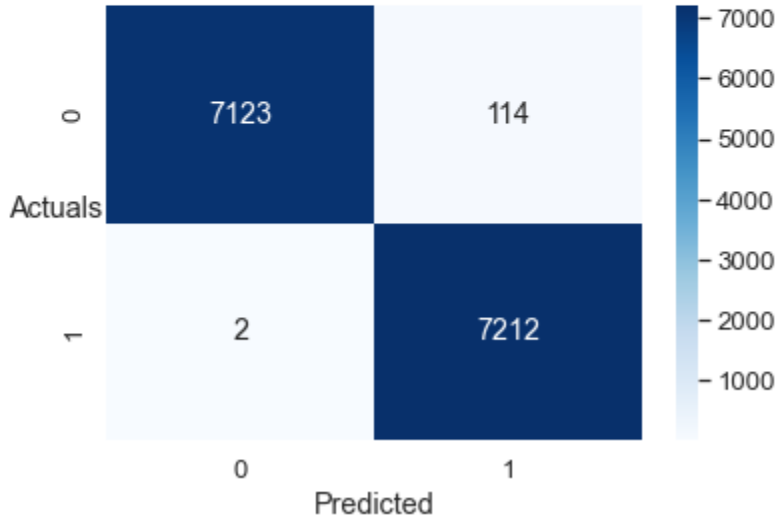
```
The model score for KNN training set is 0.9919728738495606
```

```
The model score for KNN testing set is 0.9861610849709382
```

Mobile KNN Training Set

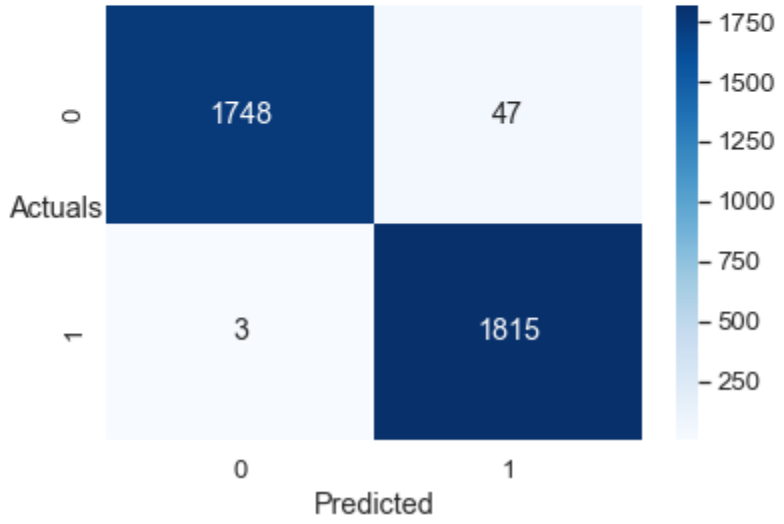
The classification report for KNN set is

	precision	recall	f1-score	support
0.0	1.00	0.98	0.99	7237
1.0	0.98	1.00	0.99	7214
accuracy			0.99	14451
macro avg	0.99	0.99	0.99	14451
weighted avg	0.99	0.99	0.99	14451



KNN Testing Set

```
The classification report for KNN testing set is
              precision    recall  f1-score   support

         0.0       1.00      0.97      0.99      1795
         1.0       0.97      1.00      0.99      1818

    accuracy                           0.99      3613
   macro avg       0.99      0.99      0.99      3613
weighted avg       0.99      0.99      0.99      3613
```

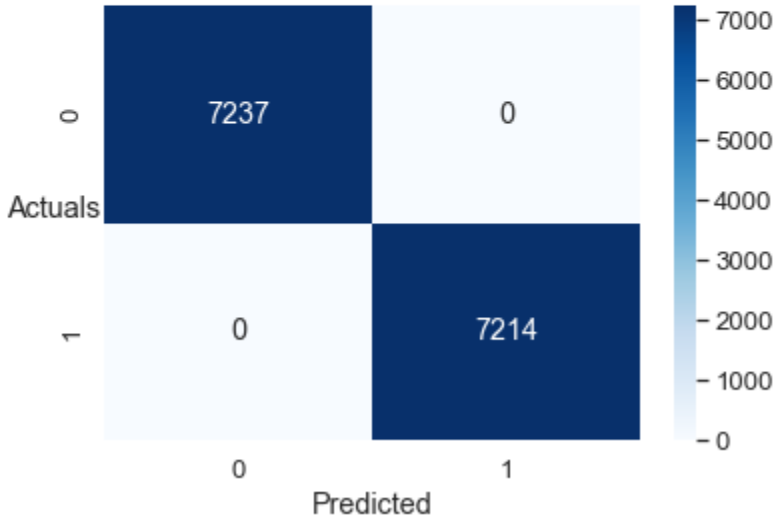


Model Tuning and business implications

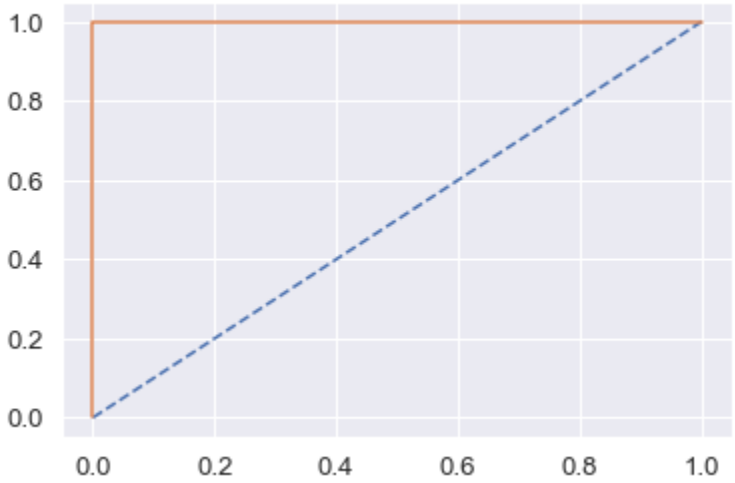
Mobile Bagging training set

```
The model score for Bagging training set is 1.0
```

```
The classification report for Bagging training set is
              precision    recall  f1-score   support

         0.0       1.00      1.00      1.00      7237
         1.0       1.00      1.00      1.00      7214

    accuracy                           1.00     14451
   macro avg       1.00      1.00      1.00     14451
weighted avg       1.00      1.00      1.00     14451
```



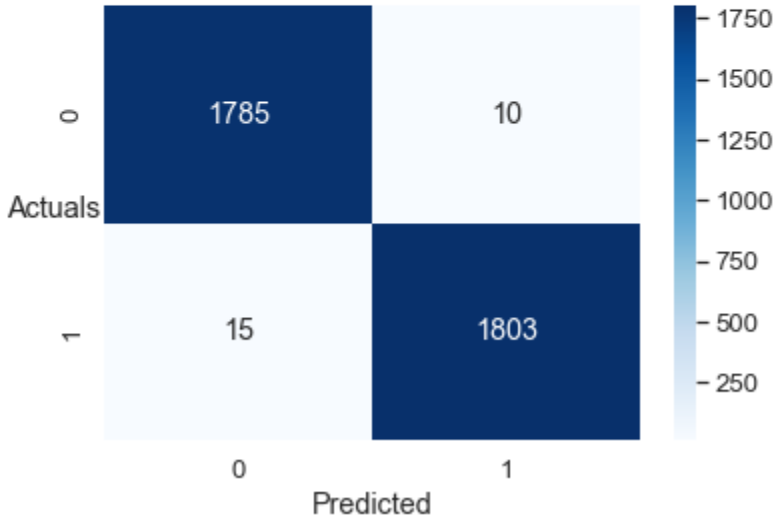

```
The AUC score for Bagging training set is: 1.000
```



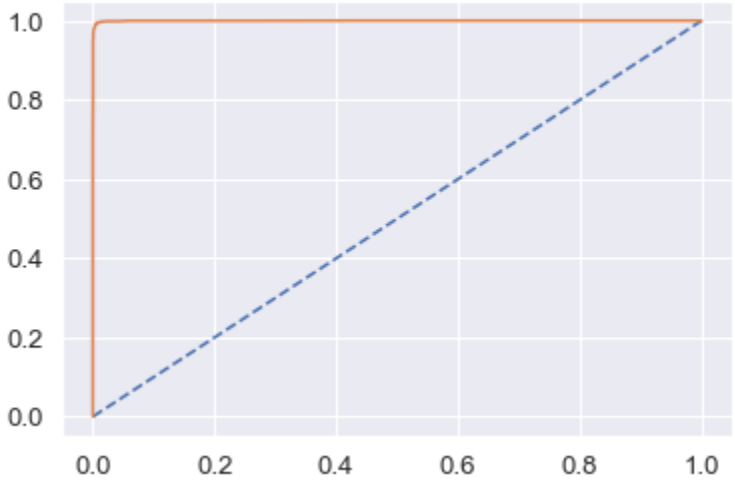
Mobile Bagging testing set

```
The model score for Bagging testing set is 0.9929889298892989
```

```
The classification report for Bagging testing set is
              precision    recall  f1-score   support

         0.0       0.99      0.99      0.99      1795
         1.0       0.99      0.99      0.99      1818

    accuracy                           0.99      3613
   macro avg       0.99      0.99      0.99      3613
weighted avg       0.99      0.99      0.99      3613
```

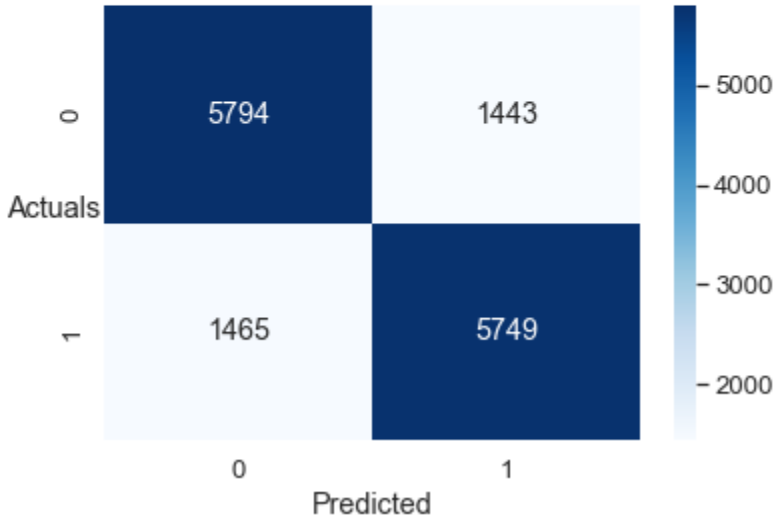



```
The AUC score for Bagging testing set is: 1.000
```

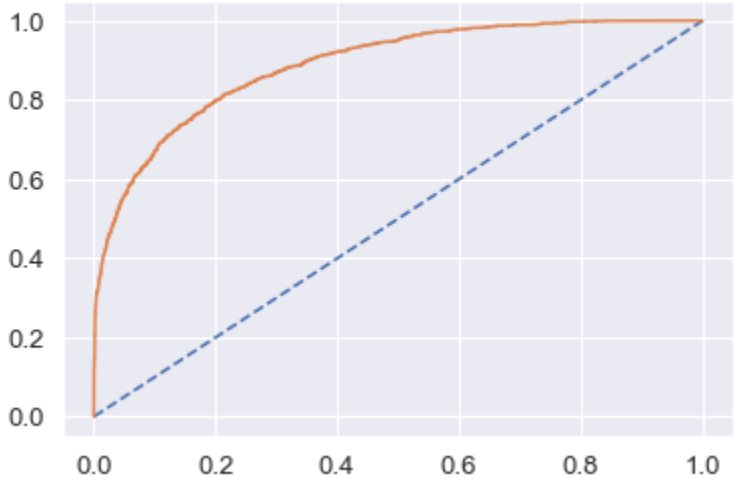


Mobile AdaBoosting Training Set

```
The model score for AdaBoosting training set is 0.7982442265105979

The classification report for Adaboosting training set is
              precision    recall  f1-score   support

         0.0       0.80      0.80      0.80      7237
         1.0       0.80      0.80      0.80      7214

    accuracy                           0.80     14451
   macro avg       0.80      0.80      0.80     14451
weighted avg       0.80      0.80      0.80     14451
```

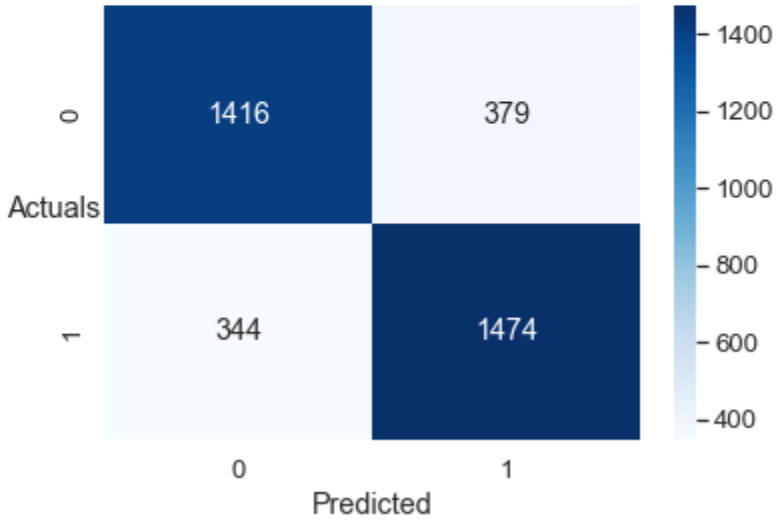



```
The AUC score for AdaBoosting training set is: 0.888
```

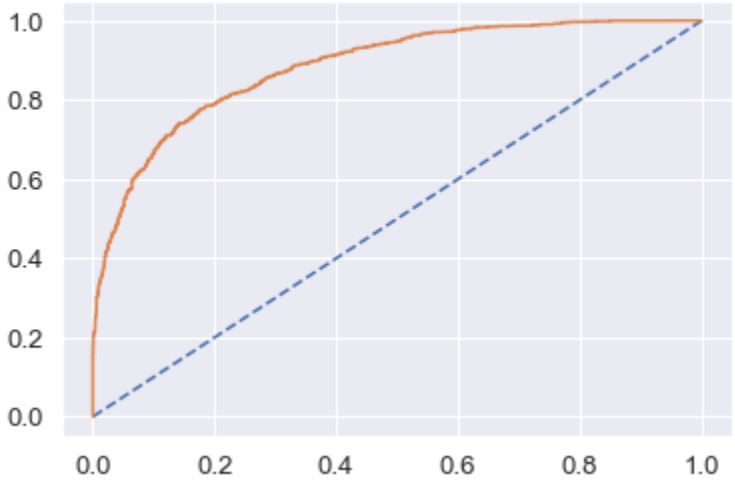


Mobile AdaBoosting Testing Set

```
The model score for AdaBoosting testing set is 0.794280442804428

The classification report for Adaboosting testing set is
              precision    recall  f1-score   support

         0.0       0.80      0.79      0.80      1795
         1.0       0.80      0.81      0.80      1818

    accuracy                           0.80      3613
   macro avg       0.80      0.80      0.80      3613
weighted avg       0.80      0.80      0.80      3613
```



```
The AUC score for AdaBoosting testing set is: 0.884
```

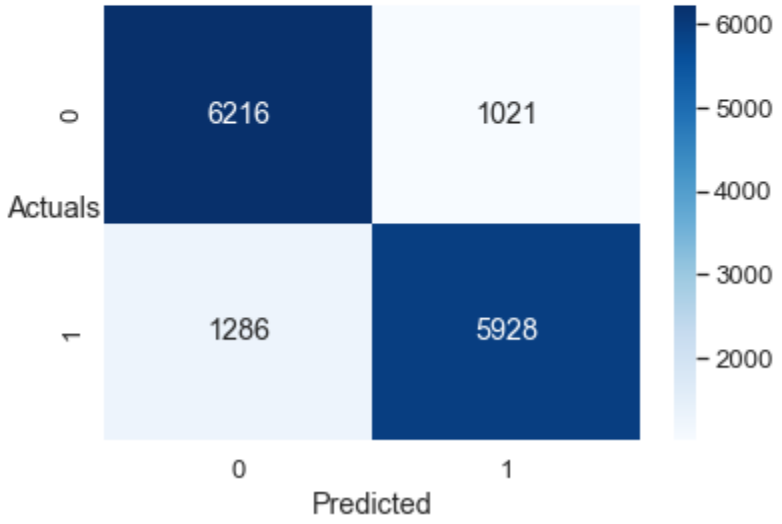


Mobile Gradient Boosting Training Set

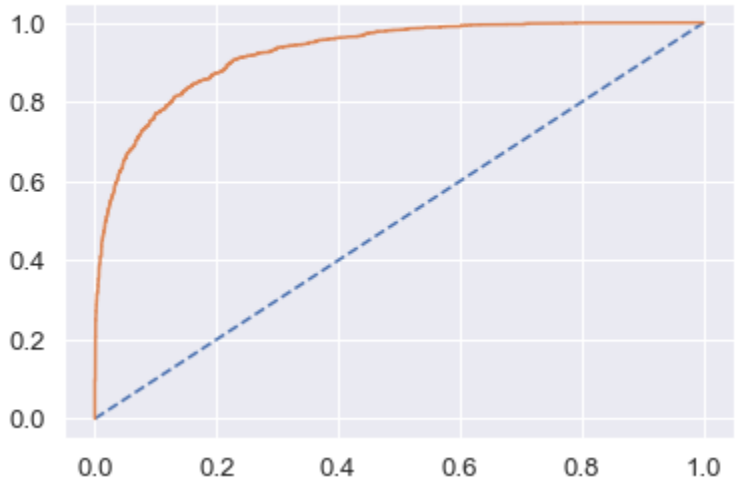
```
The model score for GradientBoosting training set is 0.8390540968048086
```

The classification report for Gradientboosting training set is

	precision	recall	f1-score	support
0.0	0.83	0.86	0.84	7237
1.0	0.85	0.82	0.84	7214
accuracy			0.84	14451
macro avg	0.84	0.84	0.84	14451
weighted avg	0.84	0.84	0.84	14451

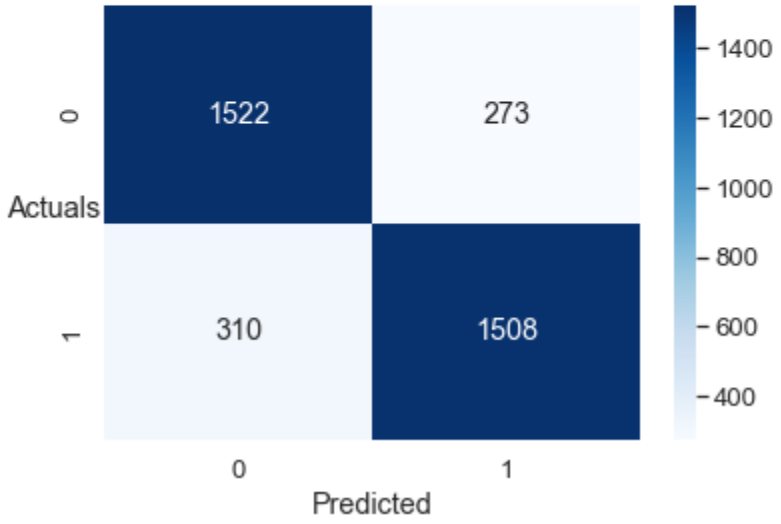


The AUC score for GradientBoosting training set is: 0.925

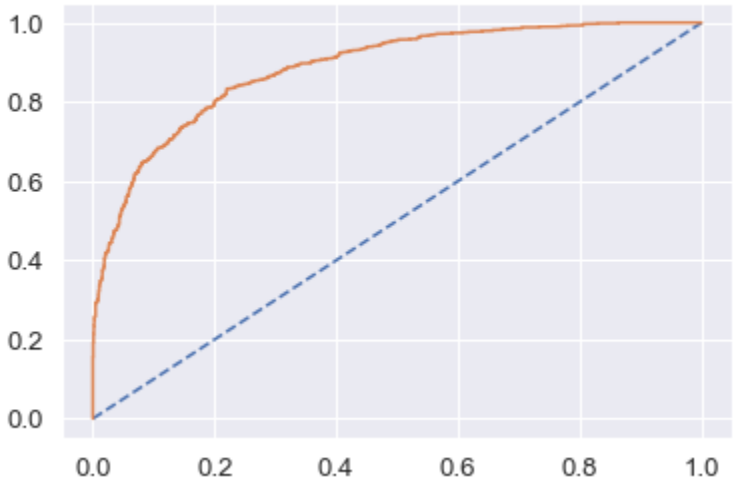


Mobile Gradient Boosting Testing Set

```
The model score for GradientBoosting testing set is 0.8363468634686347
The classification report for Gradientboosting testing set is
              precision    recall  f1-score   support

         0.0       0.83      0.85      0.84      1795
         1.0       0.85      0.83      0.84      1818

    accuracy                           0.84      3613
   macro avg       0.84      0.84      0.84      3613
weighted avg       0.84      0.84      0.84      3613
```



```
The AUC score for GradientBoosting testing set is: 0.886
```



Comparing all the models

	LR Train	LR Test	LDA Train	LDA Test	NB Train	NB Test	CART Train	CART Test	RFC Train	RFC Test	Bagging Train	Bagging Test	Ada Boosting Train	Ada Boosting Test	KNN Train	KNN Test	Gradient Boosting Train	Gradient Boosting Test
Precision	0.710	0.712	0.708	0.716	0.658	0.662	1.0	0.987	1.0	0.997	1.0	0.994	0.799	0.795	0.984	0.975	0.853	0.847
Recall	0.706	0.733	0.702	0.732	0.739	0.764	1.0	0.985	1.0	0.996	1.0	0.992	0.797	0.811	1.000	0.998	0.822	0.829
F1 Score	0.708	0.723	0.705	0.724	0.696	0.710	1.0	0.986	1.0	0.996	1.0	0.993	0.798	0.803	0.992	0.986	0.837	0.838
Accuracy	0.709	0.717	0.707	0.719	0.678	0.685	1.0	0.986	1.0	0.996	1.0	0.993	0.799	0.800	0.992	0.986	0.840	0.839
AUC Score	0.777	0.783	0.776	0.782	0.757	0.765	1.0	0.986	1.0	1.000	1.0	1.000	0.888	0.886	1.000	0.997	0.930	0.886

Inferences for Mobile devices based on Model Building

vii. Logistic Regression

The Logistic Regression has come up with poor accuracy of 70.9% on the train set and only 71.7% accuracy on test set. The precision also is not up to the mark i.e. 71.0% for train and 71.2% for the test datasets.

viii. Linear Discriminant Analysis

The LDA has also not performed well. It has given an accuracy score of 70.7% on Train set and 71.9% on the Test set. The precision also isn't up to the mark. Precision is only 70.8% for train set and 71.6% for the test set.

ix. Naïve Bayes Model

The Naïve Bayes model has also not shown very poor performance. The accuracy score is 67.8% and 68.5% on train as well as test split datasets. The

precision is poor for this model. It is 65.8% & 66.2% for train as well as test sets.

x. Decision Tree Classifier

The Decision Tree Classifier (CART) model has performed reasonably well. The accuracy score for train and test dataset is 100% and 98.6%. The precision is also very good. It is 100% and 98.7% for train and test set respectively.

xi. Random Forest Classifier

The performance of the Random Forest Classifier is the best amongst all the models. The accuracy is 100% for the train and 99.6% for the test data sets. The precision is also 100% for training dataset and 99.7% testing dataset. The number of False positives are 6 and False Negatives is 8.

xii. K- Nearest Neighbour

This model has performed fairly but not as good as Random Forest Classifier. The accuracy is 99.2% for train and 98.6% for the test dataset. The number of False positives are 47 and False Negatives are 3.

Model Tuning (Bagging & Boosting)

After applying the model tuning technique bagging to the model the performance received was good. Accuracy for train dataset was 100% and 99.3% for the test. In the case of boosting the Gradient boosting gave accuracy of 84.0% on train and 83.9% on the test datasets.

The final recommendation to business shall be move ahead with Random Forest Classifier for mobile devices. The number of False Positives and False Negatives are very high in case of K-Nearest Neighbor.

Interpretation of the most optimum model and its implication on the business

Based on the model building approach used for both the devices laptop and mobile, we can finally conclude that the Random Forest Classifier shall be the best model for predicting the likeness of a customer for buying the product through the social media campaign by the GO-GO Air company.