Minimum viable product (MVP) to churn prediction using classification algorithms

As shown in the figure below we used the logistic regression as our baseline model to begin with, and we got the following results:

Recall: (class = 'Yes', 0.90), (class = 'No', 0.50)

Precision: (class = 'Yes', 0.83), (class = 'No', 0.64)

Based on these formulas:

Recall = $TP / TP + FN$

Precision = $TP / TP + FP$

Our aim is to reduce the type 1 error (False Negative) where's the customer will leave the company and we predicted as loyal customer.

Baseline Model:

Logestic regression Model:

```
In [42]: # logestic regression model as baseline
         logreg = LogisticRegression()
         logreg.fit(x_train, y_train)

Out[42]: LogisticRegression()

In [43]: y_pred = logreg.predict(x_test)

         # classification report
         target_names = ['Yes class', 'No class']
         print(classification_report(y_test, y_pred, target_names=target_names))

                       precision    recall  f1-score   support

            Yes class       0.83      0.90      0.87      1041
             No class       0.64      0.50      0.56       368

             accuracy                           0.79      1409
            macro avg       0.74      0.70      0.71      1409
         weighted avg       0.78      0.79      0.79      1409
```

Figure 1: Baseline model