In [1]:
```
from math import sqrt
```

In [2]:
```
def indicator(TP, FP, FN, TN):
    acc = round(((TP + TN) / (TP + FN + FP + TN)),4)
    recall_positive = round((TP / (TP + FN)),4)
    recall_negative = round((TN / (FP + TN)),4)
    precision_positive = round((TP / (TP + FP)),4)
    precision_negative = round((TN / (TN + FN)),4)
    f1_score_positive = round((2*precision_positive*recall_positive / (precision_positive + recall_positive)),4)
    f1_score_negative = round((2*precision_negative*recall_negative / (precision_negative + recall_negative)),4)

    p_0 = (TN + TP) / (TN + FP + FN + TP)
    p_c = ((TN+FN)*(TN+FP) + (FN+TP)*(FP+TP)) / (TN + FP + FN + TP)**2
    kappa = round(((p_0 - p_c) / (1-p_c)),4)

    numerator = (TP * TN) - (FP * FN)
    denominator = sqrt((TP+FP) * (TP+FN) * (TN+FP) * (TN+FN))
    mcc = round((numerator / denominator),4)

    print("\n",'1.acc:',acc,",\n",'2.recall_positive:',recall_positive,",\n",'3.recall_negative:',recall_negative,
          ",\n",'4.precision_positive:',precision_positive,",\n",'5.precision_negative:',precision_negative,
          ",\n",'6.f1_score_positive:',f1_score_positive,",\n",'7.f1_score_negative:',f1_score_negative,",\n",'8.kappa
```

Acc (Accuracy) : 不適用於不平衡資料，故不以此為衡量標準
recall_positive : 實際為Positive的樣本中，被正確判斷為Positive的比率
recall_negative : 實際為Negative的樣本中，被正確判斷為Negative的比率
precision_positive : 預測為Positive樣本中，實際亦為Positive的比率
precision_negative : 預測為Negative的樣本中，實際亦為Negative的比率
f1_score : precision和recall的調和平均
kappa : Cohen's kappa係數是用來分析兩個審查者對於類別項目評分的一致性，(0,1)越高越好
mcc : Matthews correlation coefficient常用於不平衡類別衡量標準,(-1,1), 1表示完美預測, -1表示完全不一致

ModelA

In [3]:
```
TP, FP, TN, FN = (853, 341, 7230, 576)
```

In [4]:
```
indicator(TP, FP, FN, TN)
```

```
 1.acc: 0.8981 ,
 2.recall_positive: 0.5969 ,
 3.recall_negative: 0.955 ,
 4.precision_positive: 0.7144 ,
 5.precision_negative: 0.9262 ,
 6.f1_score_positive: 0.6504 ,
 7.f1_score_negative: 0.9404 ,
 8.kappa: 0.5913 ,
 9.mcc: 0.5946
```

ModelB

In [5]:
```
TP, FP, TN, FN = (846, 316, 7255, 583)
```

In [6]:
```
indicator(TP, FP, FN, TN)
```

```
 1.acc: 0.9001 ,
 2.recall_positive: 0.592 ,
 3.recall_negative: 0.9583 ,
 4.precision_positive: 0.7281 ,
 5.precision_negative: 0.9256 ,
 6.f1_score_positive: 0.653 ,
 7.f1_score_negative: 0.9417 ,
 8.kappa: 0.5954 ,
 9.mcc: 0.5998
```

由以上結果可知:
1.Acc : ModelA > ModelB , 在不平衡資料當中，此並非為衡量的主要標準
2.recall_positive : ModelA > ModelB, 表示實際為Positive的樣本中，被正確判斷為Positive的比率ModelA比ModelB高一點
3.recall_negative : ModelA < ModelB, 表示實際為Negative的樣本中，被正確判斷為Negative的比率ModelA比ModelB低一些
4.precision_positive : ModelA < ModelB, 表示預測為Positive樣本中，實際亦為Positive的比率ModelA比ModelB低
5.precision_negative : ModelA > ModelB, 表示預測為Negative的樣本中，實際亦為Negative的比率ModelA比ModelB高
6.f1_score_positive : ModelA < ModelB, 表示對於Positive的預測ModelA比ModelB低

7.f1_score_negative : ModelA < ModelB, 表示對於Negative的預測ModelA比ModelB低
8.kappa : ModelA < ModelB, 表示ModelB比ModelA好
9.mcc : ModelA < ModelB, 表示ModelB比ModelA好

綜合以上衡量標準, ModelB比ModelA好

In [ ]: