Detección de Fraude con tarjeta de crédito

Eguivar Villca Shirley Carminia

LIMPIEZA DE DATOS :

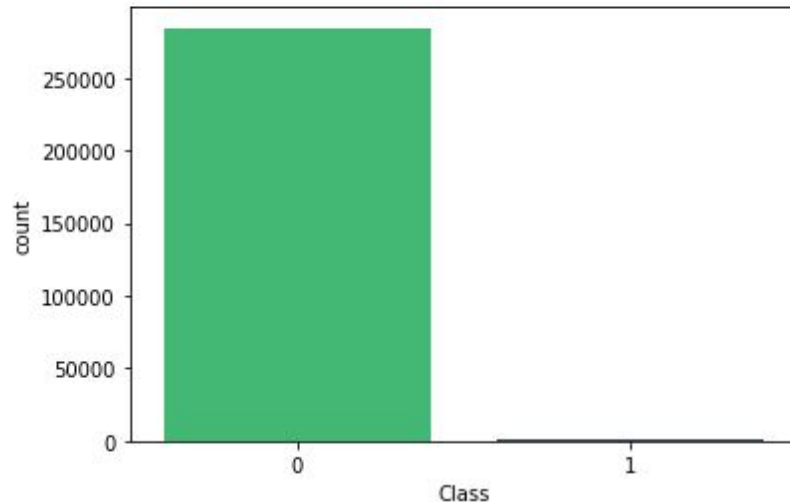Usando undersampling para balancear el dataset

DESPUES

ANTES
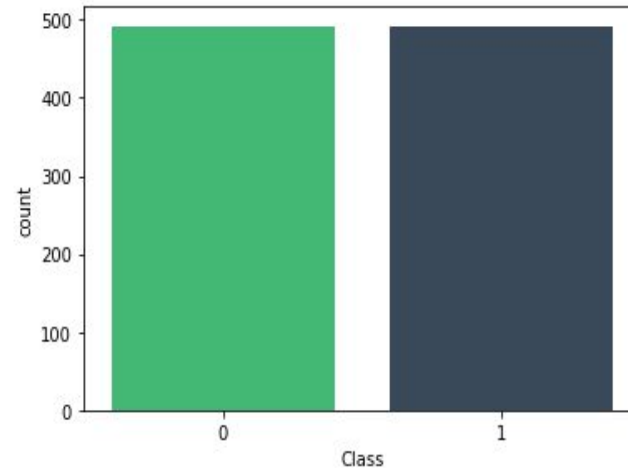
Modelos Aplicados

Regresión logística
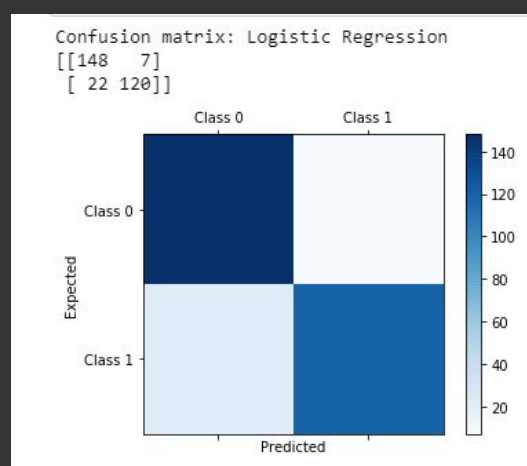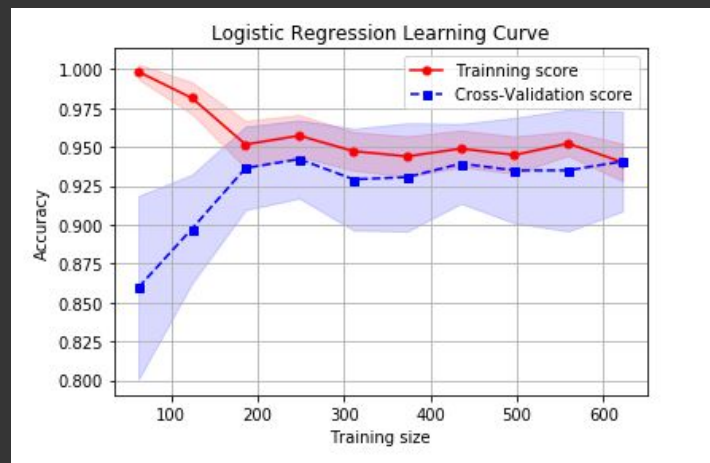
Arbol de decision

Conclusiones

El modelo de Regresión logística Tiene Mejor desempeño.

Regresión logística

```
accuracy = metrics.accuracy_score(y_pred,y_test)
print("accuracy:",accuracy)
precision = metrics.precision_score(y_pred,y_test)
print("precision:",precision)
f1_score = metrics.f1_score(y_pred,y_test)
print("f1_score:",f1_score)
recall_score = metrics.recall_score(y_pred,y_test)
print("recall_score/sensitivity:",recall_score)
```
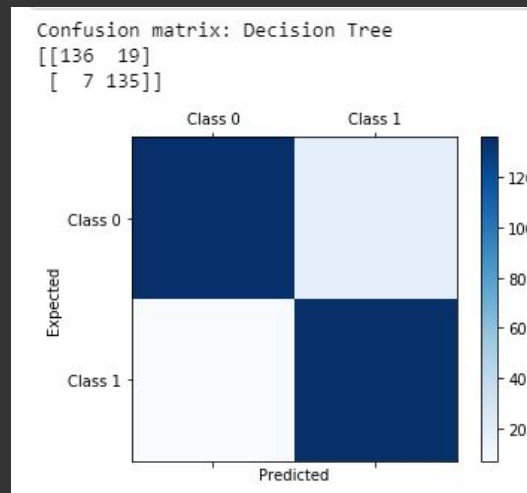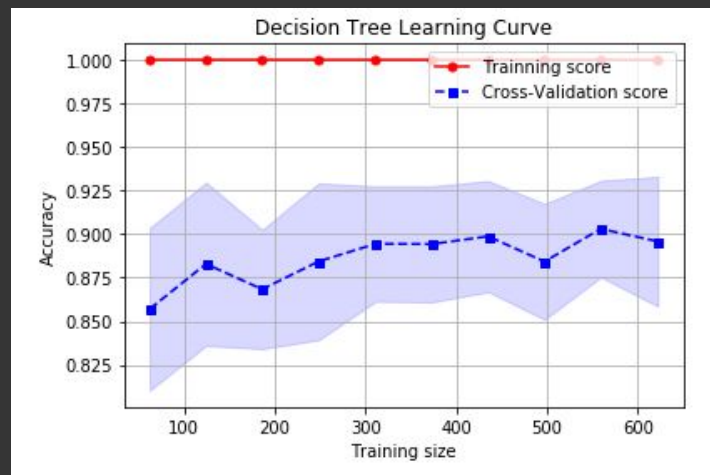
```
accuracy: 0.9528619528619529
precision: 0.9436619718309859
f1_score: 0.9503546099290779
recall_score/sensitivity: 0.9571428571428572
```

Arbol de decision

```
accuracy = metrics.accuracy_score(y_pred,y_test)
print("accuracy:",accuracy)
precision = metrics.precision_score(y_pred,y_test)
print("precision:",precision)
f1_score = metrics.f1_score(y_pred,y_test)
print("f1_score:",f1_score)
recall_score = metrics.recall_score(y_pred,y_test)
print("recall_score/sensitivity:",recall_score)
```

```
accuracy: 0.9124579124579124
precision: 0.9507042253521126
f1_score: 0.9121621621621622
recall_score/sensitivity: 0.8766233766233766
```


Algorithm Comparison