

Se realizaron análisis entre diferentes variables para determinar la correlación que tenían entre las variables categoricas:

Para el primer caso se eligió como variable dependiente “status” y como variables dependientes 'monto_financiado', 'costo_total' y 'enganche' y se realizó el análisis para obtener la precisión, exactitud, sensibilidad y el puntaje del modelo. Los resultados se obtuvieron probando entre micro, macro y weighted para obtener los mejores resultados:

Precisión del modelo:

0.7551319648093842

Exactitud del modelo:

0.7551319648093842

Sensibilidad del modelo:

0.7551319648093842

Puntaje F1 del modelo:

0.7551319648093842

Este procedimiento se realizó con distintas variables como se muestran a continuación

Caso 2.

Vars_Indep1= df2[['precio', 'costo_total', 'monto_financiado']]

Var_Dep1= df2['fraude']

Caso 3.

Vars_Indep2= df2[['costo_total', 'descuento', 'monto_financiado']]

Var_Dep2= df2['inversion']

Caso 4.

Vars_Indep4= df2[['precio', 'enganche', 'costo_total']]

Var_Dep4= df2['reautorizacion']

Caso 5. En este caso se convirtió la variable marca en una variable de tipo dicotómica para poder realizar su análisis

Vars_Indep5= df2[['precio', 'descuento', 'enganche']]

Var_Dep5= df2['marca']

Caso 6. De igual manera la variable status_cuenta se convirtió a variable dicotómica

Vars_Indep6= df2[['monto_financiado', 'costo_total', 'enganche']]

Var_Dep6= df2['status_cuenta']

Caso 7. Se convirtió en variable dicotómica

Vars_Indep7= df2[['precio', 'descuento', 'edad_cliente']]

Var_Dep7= df2['status']

Caso 8.

Vars_Indep8= df2[['precio', 'costo_total', 'limite_credito']]

```
Var_Dep8= df2['inversion']
```

Caso 9.

```
Vars_Indep9= df2[['monto_financiado', 'score_buro', 'porc_eng']]
```

```
Var_Dep9= df2['reautorizacion']
```

Caso 10.

```
Vars_Indep10= df2[['edad_cliente', 'precio', 'enganche']]
```

```
Var_Dep10= df2['fraude']
```

Como se muestra anteriormente se realizaron análisis con la misma variable dependiente, pero con distintas variables independientes. A continuación se muestra una tabla con todos los resultados obtenidos por cada caso:

	Precision	Exactitud	Sensibilidad	Puntaje F1
0	0.927126	0.755132	0.755132	0.755132
1	0.921554	0.921554	0.336610	0.921554
2	0.754985	0.754985	0.754985	0.754985
3	0.967889	0.967889	0.967889	0.967889
4	0.830645	0.830645	0.830645	0.830645
5	0.525464	0.525367	0.525367	0.525367
6	0.759971	0.759971	0.759971	0.759971
7	0.216569	0.216569	0.216569	0.216569
8	0.963636	0.963636	0.963636	0.963636
9	0.927126	0.927126	0.927126	0.927126

Se observa que “status” se comparo primero con 'monto_financiado', 'costo_total' y 'enganche' obteniendo como resultado:

0	0.927126	0.755132	0.755132	0.755132
---	----------	----------	----------	----------

y despues con 'precio', 'descuento' y 'edad_cliente' obteniendo:

6	0.759971	0.759971	0.759971	0.759971
---	----------	----------	----------	----------

Por lo que se observa que la precision del caso uno fue mayor, pero la exactitud, sensibilidad y el puntaje fueron mayores en el caso 7, por lo que el caso 7 tiene una mejor correlación con el “status”

Caso 2 vs Caso 10

En este caso se analizó la variable “fraude” contra 'precio', 'costo_total', 'monto_financiado'(caso 2)

1	0.921554	0.921554	0.336610	0.921554
---	----------	----------	----------	----------

VS 'edad_cliente', 'precio', 'enganche'(caso10)

9	0.927126	0.927126	0.927126	0.927126
---	----------	----------	----------	----------

En este caso se observa que ambas tuvieron buenos resultados, pero en cuanto a la sensibilidad del modelo de caso 2, fue muy baja, además de que en el caso 10 los valores son ligeramente más alto, representando una mejor correlación.

Caso 3 VS Caso 8

En este caso se analizó la variable “inversión” contra 'costo_total', 'descuento' y “monto_financiado”(Caso 3)

2	0.754985	0.754985	0.754985	0.754985
---	----------	----------	----------	----------

VS 'precio', 'costo_total', 'limite_credito' (Caso 8)

7	0.216569	0.216569	0.216569	0.216569
---	----------	----------	----------	----------

Se observa que los resultados del caso 3 fueron mejores en comparación al caso 8, por lo que hay una mejor correlación entre esas variables

Caso 4 VS Caso 9

En este caso se analizó la variable “reautorizacion” contra 'precio', 'enganche', 'costo_total'(Caso 4)

3	0.967889	0.967889	0.967889	0.967889
---	----------	----------	----------	----------

VS 'monto_financiado', 'score_buro', 'porc_eng' (Caso 9)

8	0.963636	0.963636	0.963636	0.963636
---	----------	----------	----------	----------

En este caso, nuevamente los resultado obtenidos son muy similares , ambos tienen una muy buena correlación ya que se encuentra arriba del 95%, pero en este caso se correlacionaron mejor las variables del caso 9