

Reporte Comparativo – Actividad 7

Diego Garza Solís, A01705984

Primeramente, se realizó una regresión logística con todas las variables independientes, para observar cómo predirían la variable objetivo. El resultado de precisión, como es de esperarse, resulta ser de los más altos.

	Variable	Precisión	Exactitud	Puntaje F1
0	Todas	0.852553	0.796361	0.772952
1	precio	0.533682	0.514112	0.439300
2	costo_total	0.572618	0.555167	0.535671
3	enganche	0.562414	0.553767	0.562543
4	monto_financiado	0.559116	0.540005	0.506507
5	monto_accesorios	0.000000	0.490086	0.000000
6	pagos_realizados	0.538509	0.519244	0.458913
7	reautorizacion	0.000000	0.490086	0.000000
8	limite_credito	0.873582	0.812690	0.801090
9	descuento	0.526974	0.506181	0.389033
10	edad_cliente	0.542530	0.526475	0.495025

Posteriormente, se realizaron análisis entre variables seleccionadas, donde se encontró, usando el valor de F1, que la variable con mayor puntaje respecto de la variable objetivo es “limite_credito”. La mayoría del resto de las variables tienen valores modestos, cercanos a 0.50. Hay dos variables que tienen valores de 0, de los cuales se puede abstraer que el modelo solo pudo predecir una clase para la variable dicotómica creada. Las correlaciones, en general, si bien no son fuertes, son considerables, teniendo, en conjunto, sentido hacia el análisis inicial realizado con todas ellas. Tener en el archivo, además, las matrices de confusión, nos ayuda a esclarecer mejor el proceso del mecanismo de regresión logística.