

MODELO DE ORIGINACIÓN DE CRÉDITOS

Predicción del comportamiento de
pago de nuevos créditos

BANK FIVE INTERNATIONAL

EQUIPO DE TRABAJO



ALEXIS TAPIA

Líder de Equipo



ISABEL PILAR

Analista de
Datos



BARBARA HERRERA

Analista de
Datos



CLAUDIA VILLEGAS

Ingeniero de
Modelamiento



MIGUEL PEÑA

Líder de Negocio/
Ingeniero de
Modelamiento



CONTENIDO

01

DESCRIPCIÓN DEL PROBLEMA

02

OBJETIVO

03

PROCESAMIENTO DE DATOS

04

MODELOS Y RESULTADOS

05

IMPLEMENTACIÓN

06

CONCLUSIONES

DESCRIPCIÓN DEL PROBLEMA

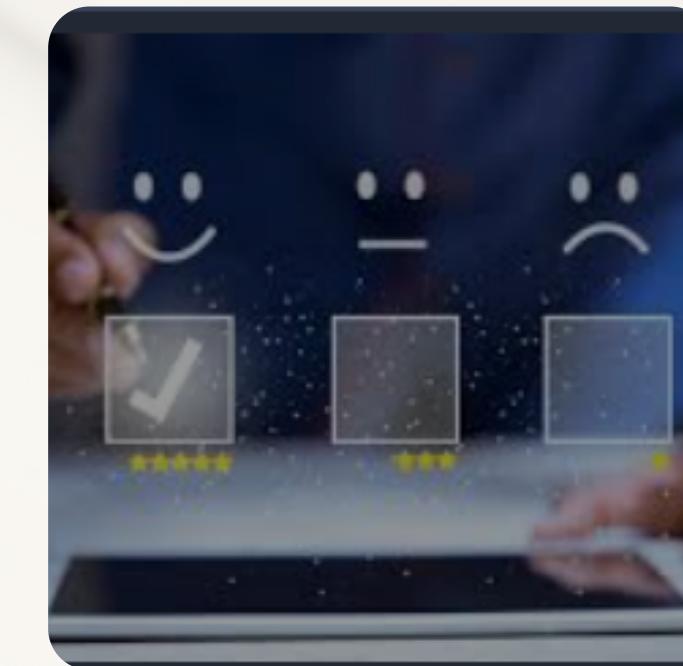


SOLICITUD DE CRÉDITO



EVALUACIÓN DE CRÉDITO

CLIENTE MOROSO



CLIENTE BUEN PAGADOR

COMPORTAMIENTO DE PAGO

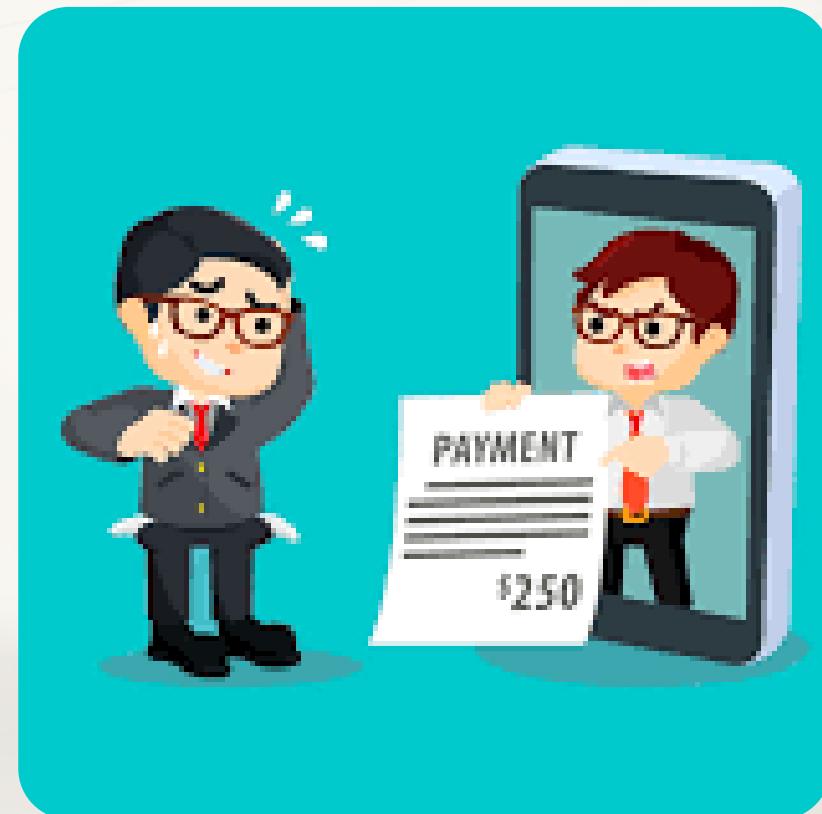
OBJETIVO

**REALIZAR UN MODELO PARA
PREDECIR TEMPRANAMENTE EL
COMPORTAMIENTO DE PAGO**

Para predecir tempranamente el comportamiento de futuros clientes respecto a si estos son buenos pagadores para el otorgamiento de un crédito de consumo, utilizando técnicas de machine learning.



VARIABLE A MODELAR

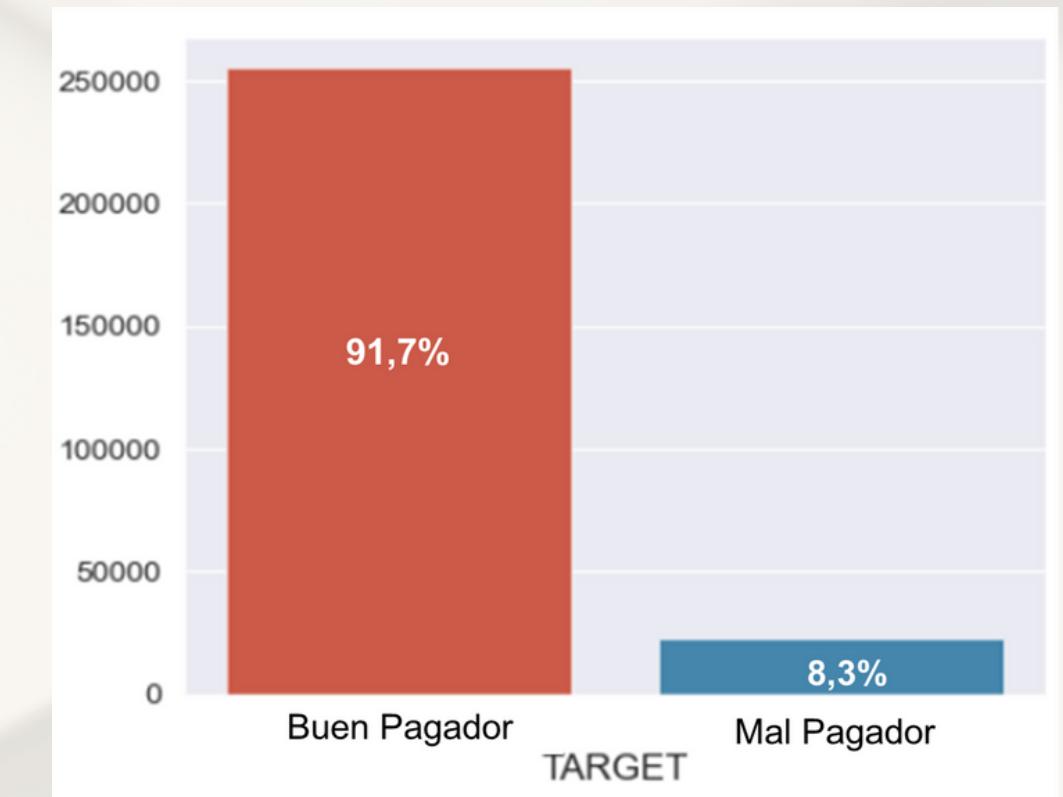


TARGET

8.3% de los clientes
presenta dificultades
para pagar

Pago atrasado por más de X
días en los primeros meses de
pago

Distribución variable objetivo

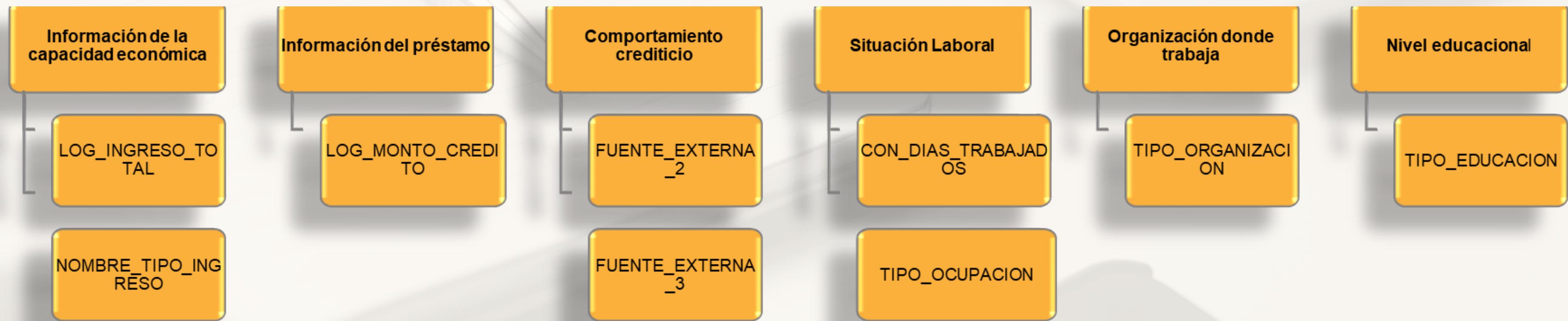


Total de registros en estudio: 278.232

FLUJO DE TRABAJO



VARIABLES DEL MODELO



VARIABLES DEL MODELO

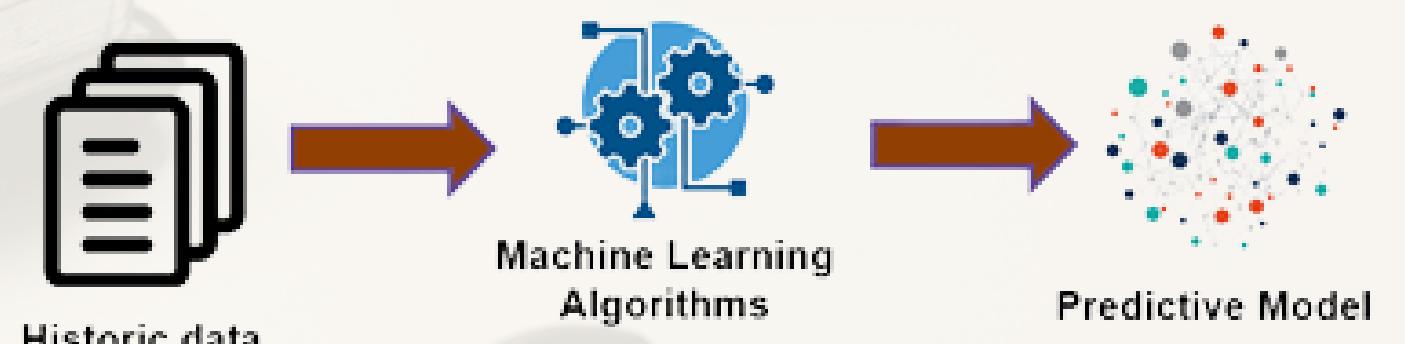


MODELOS



MODELOS DE CLASIFICACIÓN

- LogisticRegression
- LogisticGAM
- LinearDiscriminantAnalysis
- RandomForest
- GradientBoostingClassifier
- AdaBoostClassifier



RESULTADOS

Modelo	accuracy	precision	recall	f1-score	AUC	KS	GINI
Gradient Boosting	0.68	0.16	0.67	0.26	0.7383	0.3592	0.4765
GAM	0.68	0.16	0.67	0.26	0.7356	0.3559	0.4711
Regresion Logística	0.68	0.16	0.66	0.26	0.7302	0.3468	0.4604
LinearDiscriminantAnalysis	0.68	0.16	0.66	0.26	0.7301	0.3467	0.4602
AdaBoostClassifier	0.68	0.16	0.66	0.25	0.7277	0.3387	0.4553
Random Forest	0.68	0.16	0.65	0.25	0.7213	0.3336	0.4427

Tabla referencia indicadores

Poder Predictivo	Tramo AUC	Tramo KS	Tramo GINI
Excelente	>=80%	>=40%	>=60%
Bueno	>=70% y <80%	>=30% y <40%	>=40% y <60%
Regular	>=65% y <70%	>=20% y <30%	>=30% y <40%
Débil	<65%	<20%	<30%

IMPORTANCIA DE LAS VARIABLES

Variables más relevantes



EVALUACIÓN DE PUNTAJE DE CORTE

tramo_prob	n	n_buenos	n_malos	tasa_malos (%)	porc_buenos_acum (%)	porc_malos_acum (%)	dif_porc_acum (%)	utilidad_perdida_u (\$)
(0.78, 1.0]	4038	2790	1248	30.91	3.65	18.00	14.35	-175,864
(0.71, 0.78]	4502	3618	884	19.64	8.37	30.76	22.39	-77,136
(0.66, 0.71]	4004	3391	613	15.31	12.80	39.60	26.80	-39,241
(0.61, 0.66]	4469	3853	616	13.78	17.84	48.49	30.65	-25,874
(0.57, 0.61]	3950	3482	468	11.85	22.39	55.24	32.85	-8,918

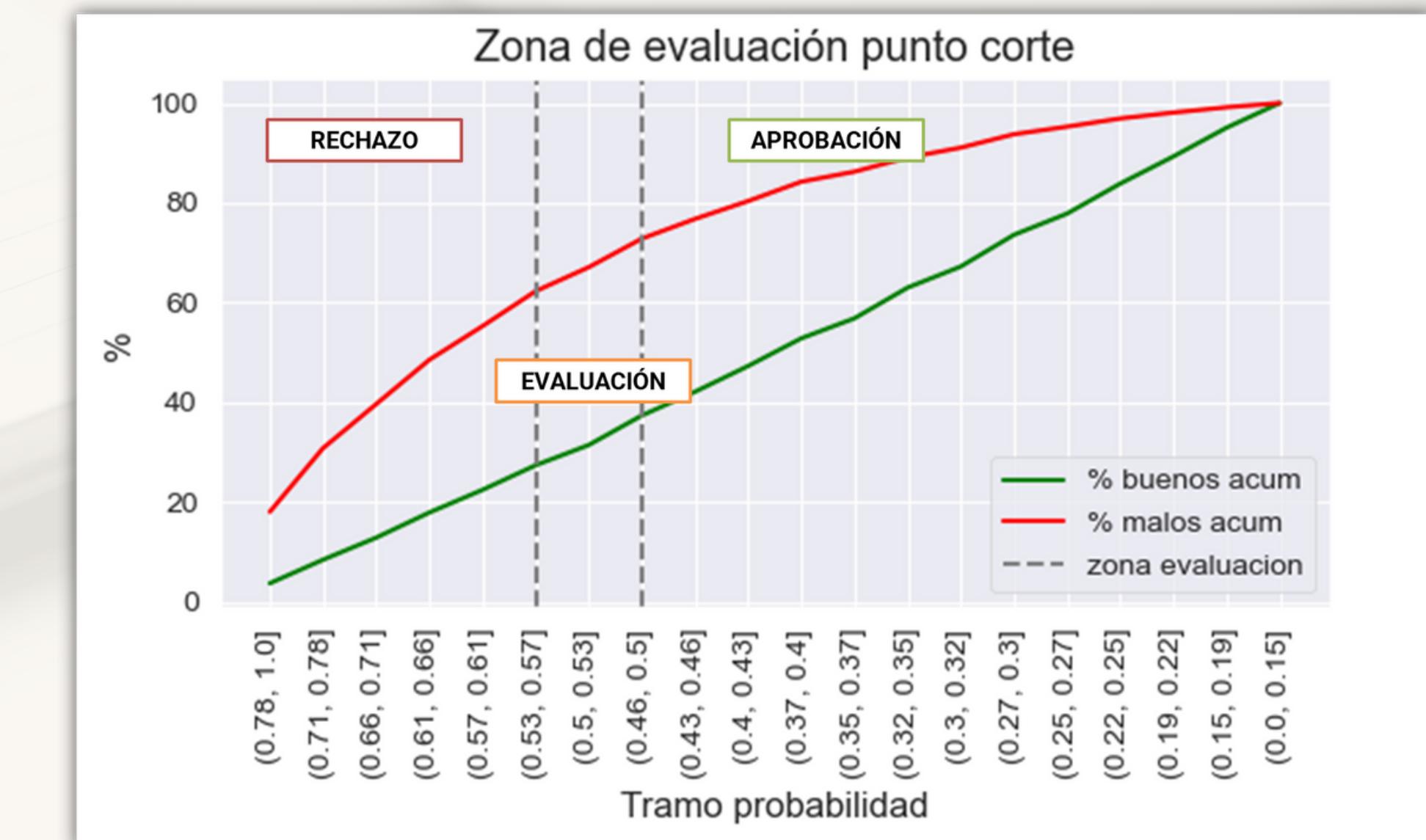
tramo_prob	n	n_buenos	n_malos	tasa_malos (%)	porc_buenos_acum (%)	porc_malos_acum (%)	dif_porc_acum (%)	utilidad_perdida_u (\$)
(0.53, 0.57]	4273	3786	487	11.40	27.33	62.26	34.93	-4,967
(0.5, 0.53]	3442	3109	333	9.67	31.39	67.07	35.68	10,122

tramo_prob	n	n_buenos	n_malos	tasa_malos (%)	porc_buenos_acum (%)	porc_malos_acum (%)	dif_porc_acum (%)	utilidad_perdida_u (\$)
(0.46, 0.5]	4935	4537	398	8.06	37.32	72.81	35.49	24,223
(0.43, 0.46]	3918	3644	274	6.99	42.08	76.76	34.68	33,609
(0.4, 0.43]	4202	3953	249	5.93	47.25	80.35	33.10	42,961
(0.37, 0.4]	4520	4251	269	5.95	52.80	84.23	31.43	42,737
(0.35, 0.37]	3204	3065	139	4.34	56.81	86.24	29.43	56,866
(0.32, 0.35]	4914	4718	196	3.99	62.97	89.07	26.10	59,930
(0.3, 0.32]	3376	3238	138	4.09	67.20	91.06	23.86	59,062
(0.27, 0.3]	5057	4870	187	3.70	73.56	93.75	20.19	62,477
(0.25, 0.27]	3360	3255	105	3.12	77.82	95.27	17.45	67,495
(0.22, 0.25]	4714	4600	114	2.42	83.83	96.91	13.08	73,685
(0.19, 0.22]	4254	4172	82	1.93	89.28	98.10	8.82	77,984
(0.15, 0.19]	4494	4420	74	1.65	95.05	99.16	4.11	80,445
(0.0, 0.15]	3844	3786	58	1.51	100.00	100.00	0.00	81,652

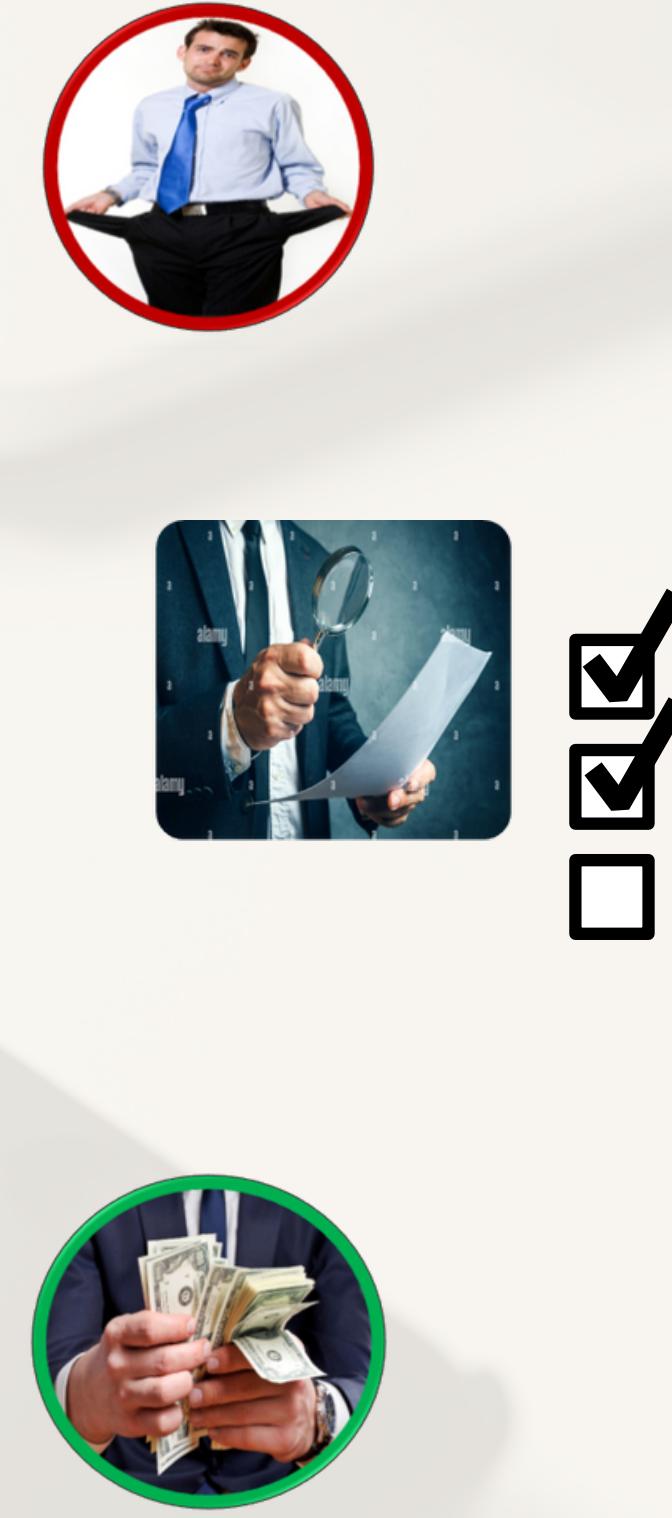
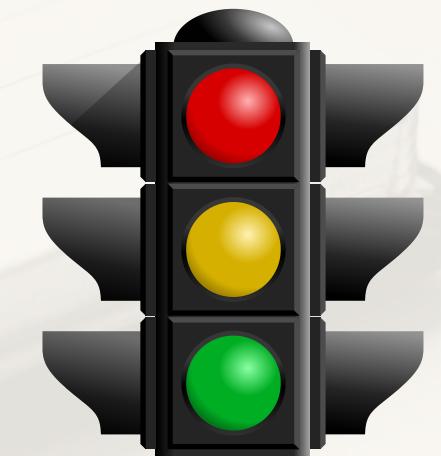
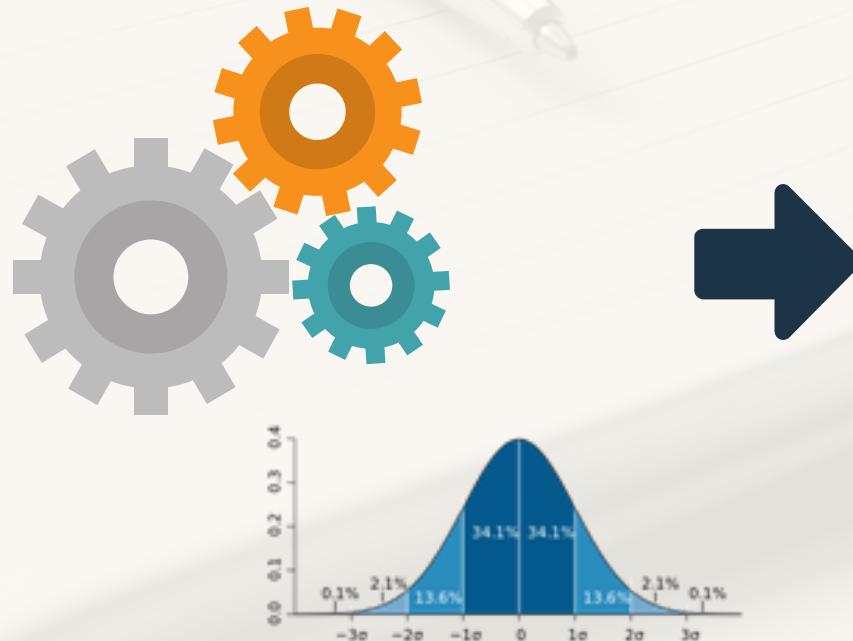
PROPUESTA USO DEL MODELO

Se proponen 3 zonas a partir del resultado del modelo.

- Si probabilidad es $> 0,57$ se debe **rechazar** el crédito.
- Si probabilidad se encuentra entre $0,5$ y $0,57$ se debe realizar una **segunda evaluación** más exhaustiva.
- Si probabilidad de ser mal pagador es $< 0,5$ se **aprueba** el crédito.



IMPLEMENTACIÓN

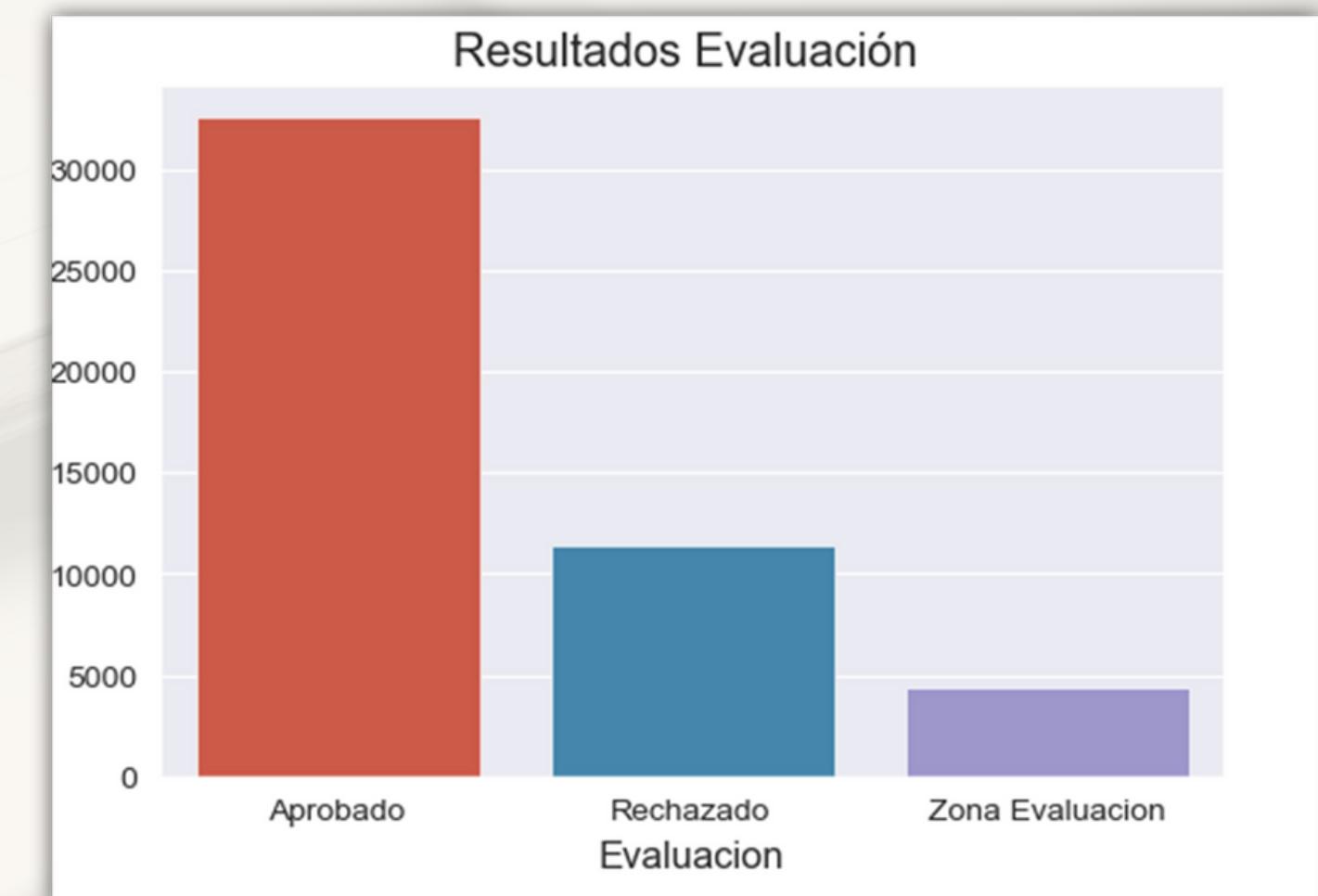


IMPLEMENTACIÓN

Se aplica modelo seleccionado a una nueva base con 48.305 créditos de consumo.

Con esto:

- Se espera que 32.584 (67.45%) sean buenos pagadores (se les apruebe el crédito)
- 4.365 (9.03%) casos deben entrar en zona de evaluación.
- 11.365 (23.51%) serían créditos rechazados.



CONCLUSIONES

- Si es posible predecir el comportamiento de pago de un cliente para determinar si es que este va a ser un buen pagador o mal pagador, mediante el uso de un modelo de machine learning.
- El modelo propuesto ofrece una alternativa de apoyo al proceso de evaluación crediticia.
- La propuesta de solución ofrece una alternativa de evaluación manual, que permita controlar posibles errores predictivos del modelo.



GRACIAS