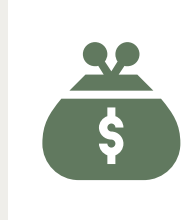




Del Dato al Negocio: Modelos Predictivos para Impulsar el Crecimiento Bancario

Cómo anticipar al cliente potencia el negocio.



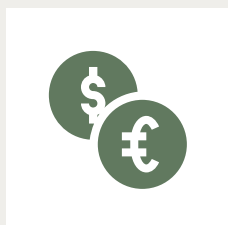
Introducción

En un contexto altamente competitivo, donde los márgenes bancarios dependen cada vez más de la capacidad de anticipar el comportamiento de los clientes, este proyecto surge como una respuesta estratégica para transformar datos históricos en decisiones comerciales inteligentes.

Mediante el desarrollo de modelos avanzados de machine learning, se ha buscado predecir qué clientes tienen mayor probabilidad de adquirir productos financieros clave como préstamos personales o tarjetas de crédito.

El objetivo final es claro:

- Optimizar los esfuerzos de marketing directo,
- Reducir costos operativos,
- Y maximizar los ingresos, captando oportunamente a los clientes con el mayor potencial.



Este proyecto no solo ofrece un análisis técnico sólido, sino que traza una hoja de ruta práctica para implementar los modelos desarrollados en la toma de decisiones comerciales del banco. Para ello, se desarrollaron y compararon dos modelos predictivos:

- LightGBM (modelo de boosting)
- Random Forest (modelo de bagging)

Ambos entrenados sobre datos históricos de transacciones de clientes, con balanceo de clases para asegurar que los modelos aprendieran adecuadamente a identificar el segmento minoritario (clientes con alta probabilidad de adquirir un producto).



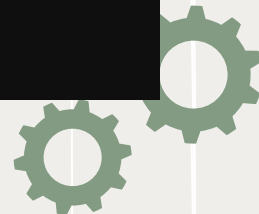
Matrices de confusión de ambos modelos

LightGMB

Evaluación del modelo LightGBM:				
	precision	recall	f1-score	support
0	0.99	0.50	0.67	35980
1	0.18	0.94	0.30	4020
accuracy			0.55	40000
macro avg	0.58	0.72	0.48	40000
weighted avg	0.91	0.55	0.63	40000
ROC AUC Score: 0.8725611450806002				

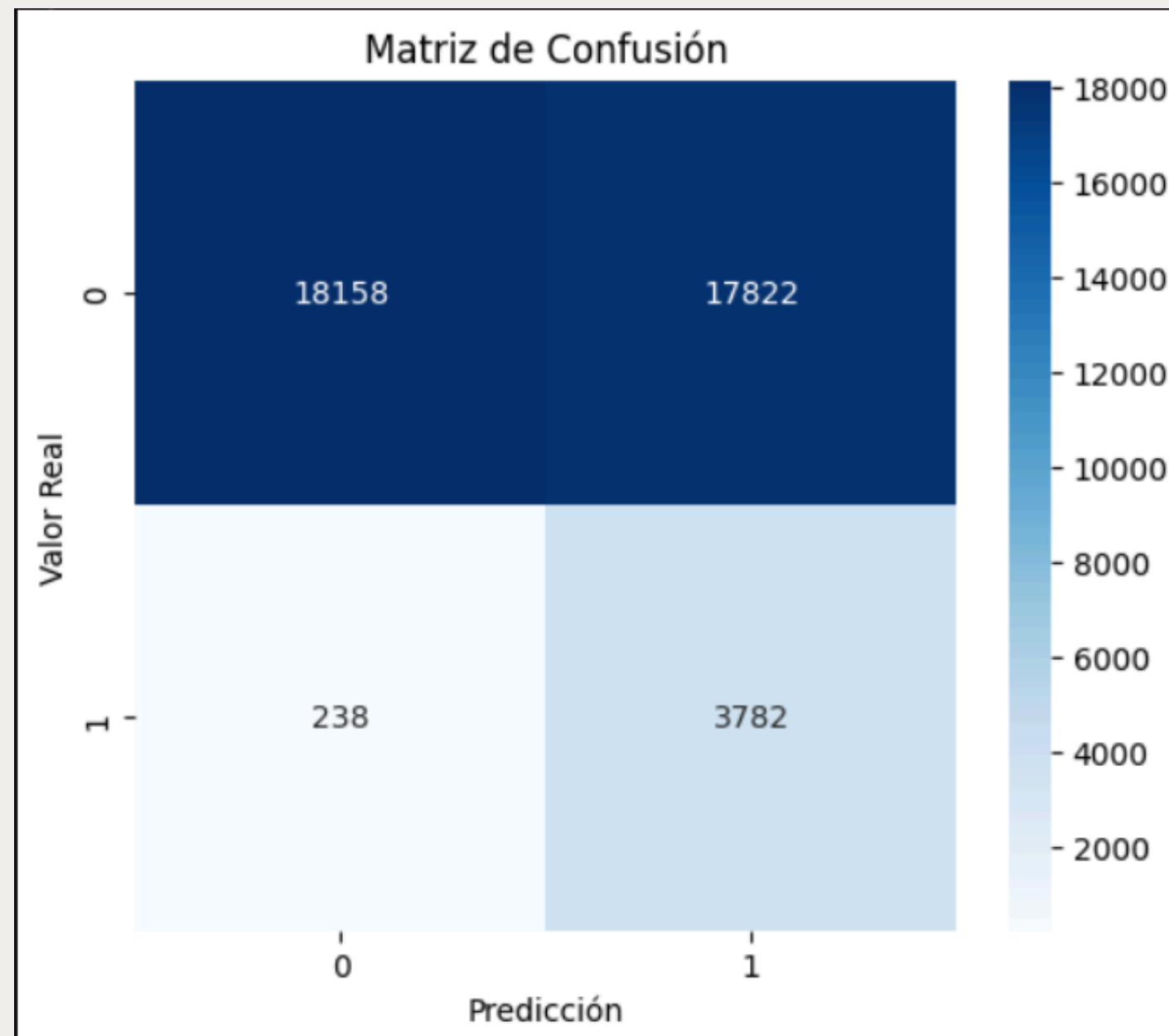
RandomForest

Evaluación del Modelo Random Forest				
AUC en el conjunto de validación (Random Forest): 0.8001				
Reporte de Clasificación en el conjunto de validación (usando umbral 0.5):				
	precision	recall	f1-score	support
0	0.94	0.86	0.90	35980
1	0.30	0.55	0.39	4020
accuracy			0.83	40000
macro avg	0.62	0.70	0.65	40000
weighted avg	0.88	0.83	0.85	40000

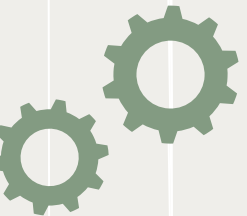
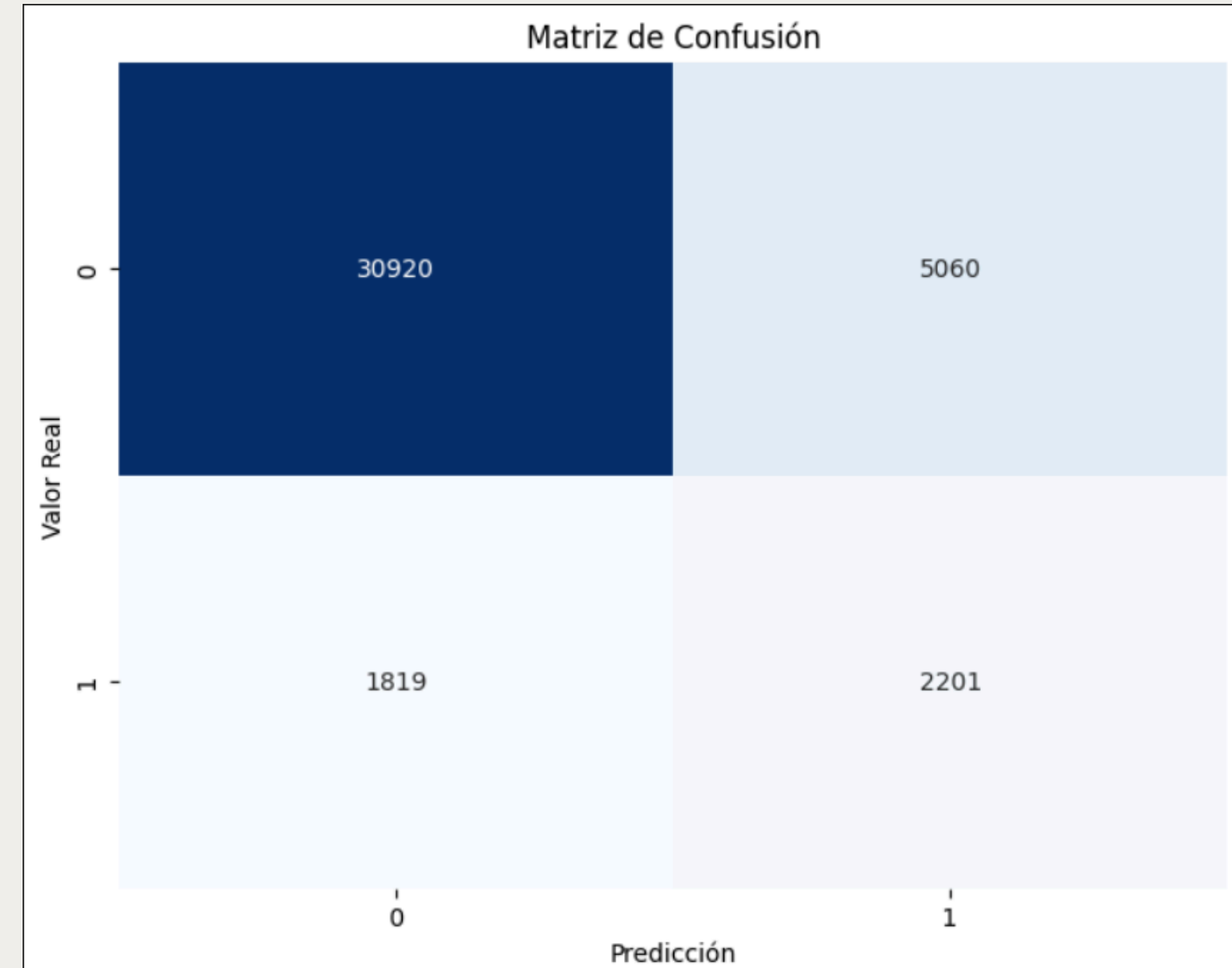


Matrices de confusión de ambos modelos

LightGMB



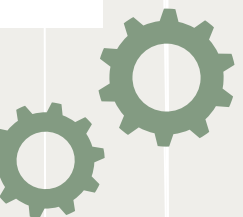
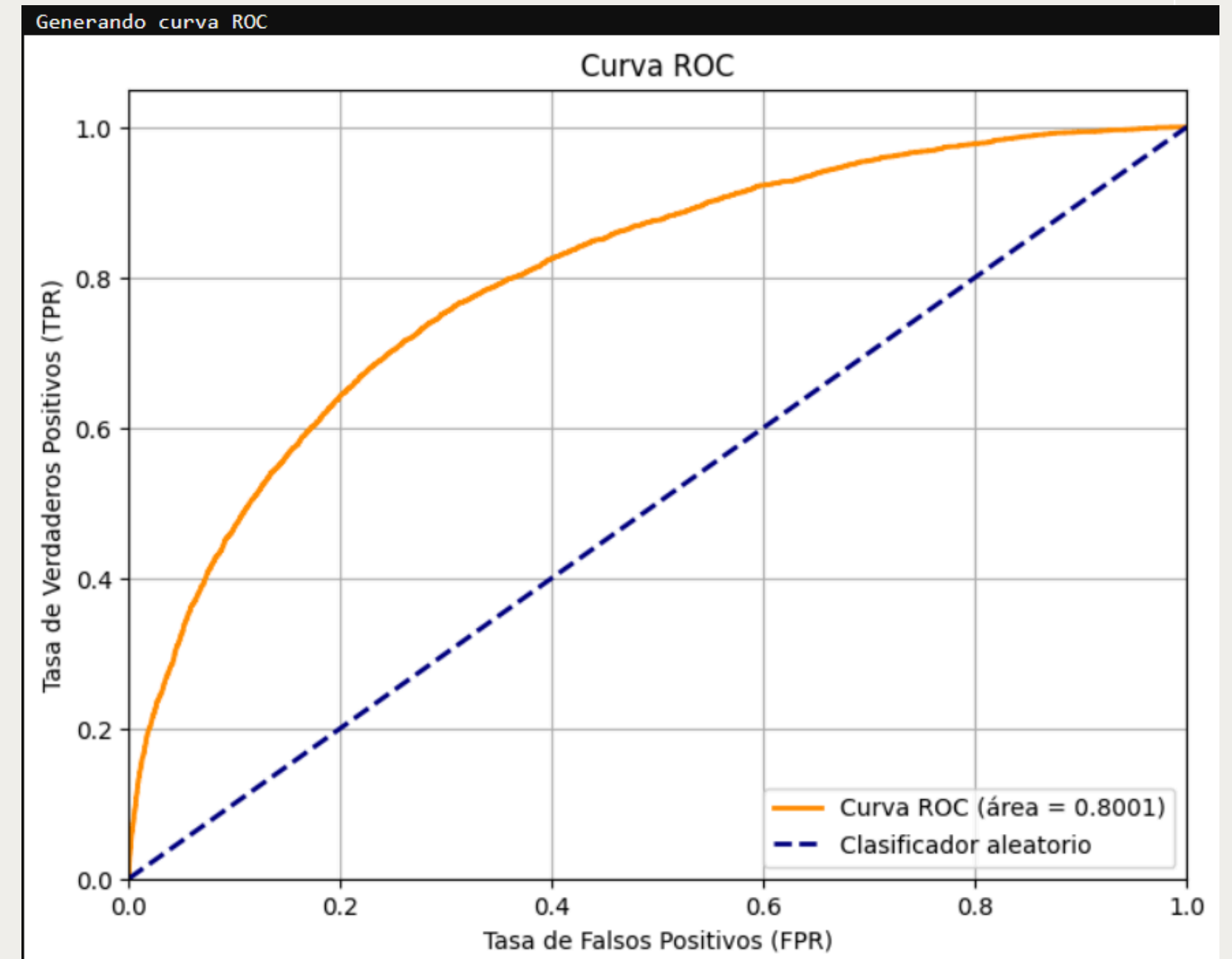
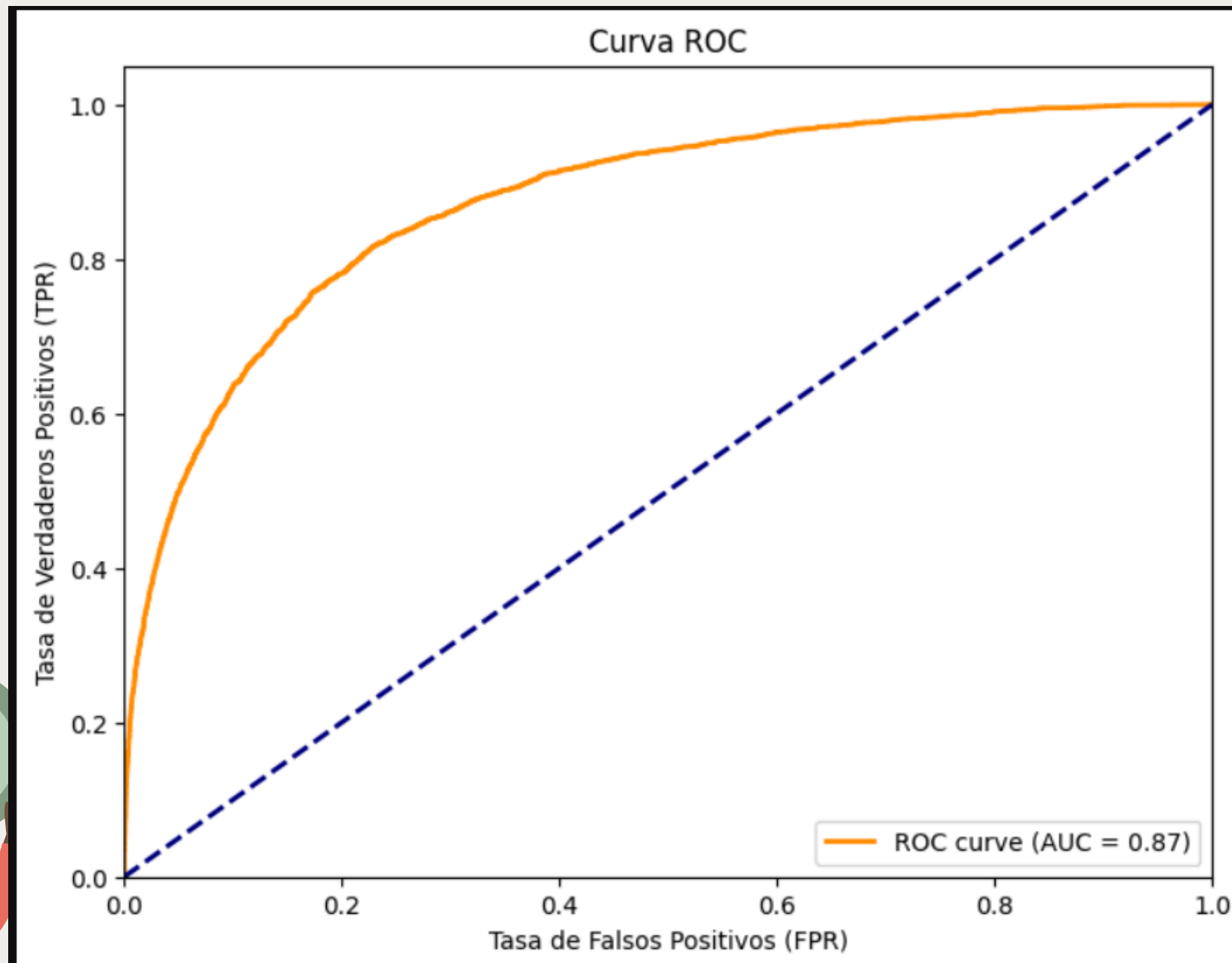
RandomForest



Curva ROC de ambos modelos

LightGMB

RandomForest





Supuestos del escenario para la simulación

Variable	Valor	Descripción
Costo contactar cliente	<u>\$5</u>	Llamada, email, gestión CRM, personal, etc.
Ganancia por cliente aceptado	<u>\$100</u>	Beneficio neto esperado si contrata préstamo o tarjeta.
Clientes en conjunto validación	<u>40,000</u>	El dataset validation.
Clientes realmente interesados (target=1)	<u>4,020 (aprox 10%)</u>	El dataset validation.





Datos clave de los modelos

Se evaluaron múltiples métricas, pero dado nuestro objetivo de negocio (maximizar clientes captados), nos centramos principalmente en:

- **Recall (sensibilidad):** para captar el mayor porcentaje posible de clientes realmente interesados.
- **Precision:** % de predicciones positivas que eran correctas (muchos falsos positivos en este caso)

Métrica	LightGBM	Random Forest
Recall (Clase 1)	0.94	0.55
Precisión (Clase 1)	0.18	0.30





Simulación costo - beneficio

Esto depende de la cantidad de positivos predichos por cada modelo. Podemos aproximarlos usando la siguiente fórmula:

- **Nº clientes contactados $\approx TP + FP = TP / Recall$**

TP= true positives, FP= False positives

Modelo	TP (Captados reales)	Recall	Estimado Contactados
LightGMB	$0.94 * 4,020 = 3,779$	0.94	$\approx 20,994$ (porque $TP / Recall = 3,779 / 0.18$)
RandomForest	$0.55 * 4,020 = 2,211$	0.55	$\approx 7,370$ ($2,211 / 0.30$)



Costo total de la campaña



Modelo	Contactados	Costo total
LightGMB	20,994	\$104,970
RandomForest	7,370	\$36,850

Ganancia total

Clientes captados por beneficio

Modelo	Captados reales (TP)	Ganancia Total
LightGMB	3,779	\$377,900
RandomForest	2,211	\$221,100



Rentabilidad Final

Beneficio neto

Modelo	Ganacia - Costos	Retorno
LightGMB	\$272,930	Más rentable sin duda alguna
RandomForest	\$184,250	Menos rentable

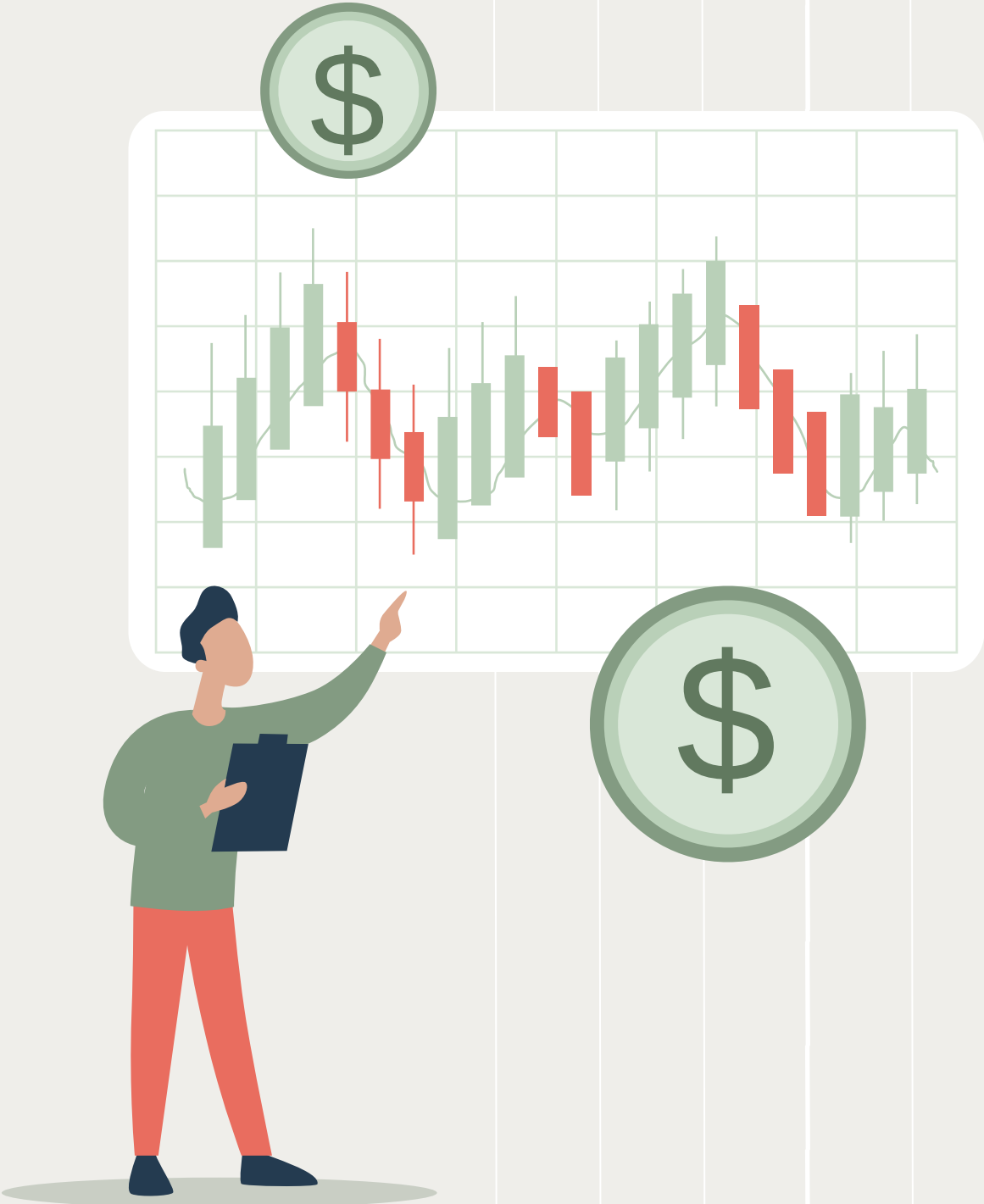
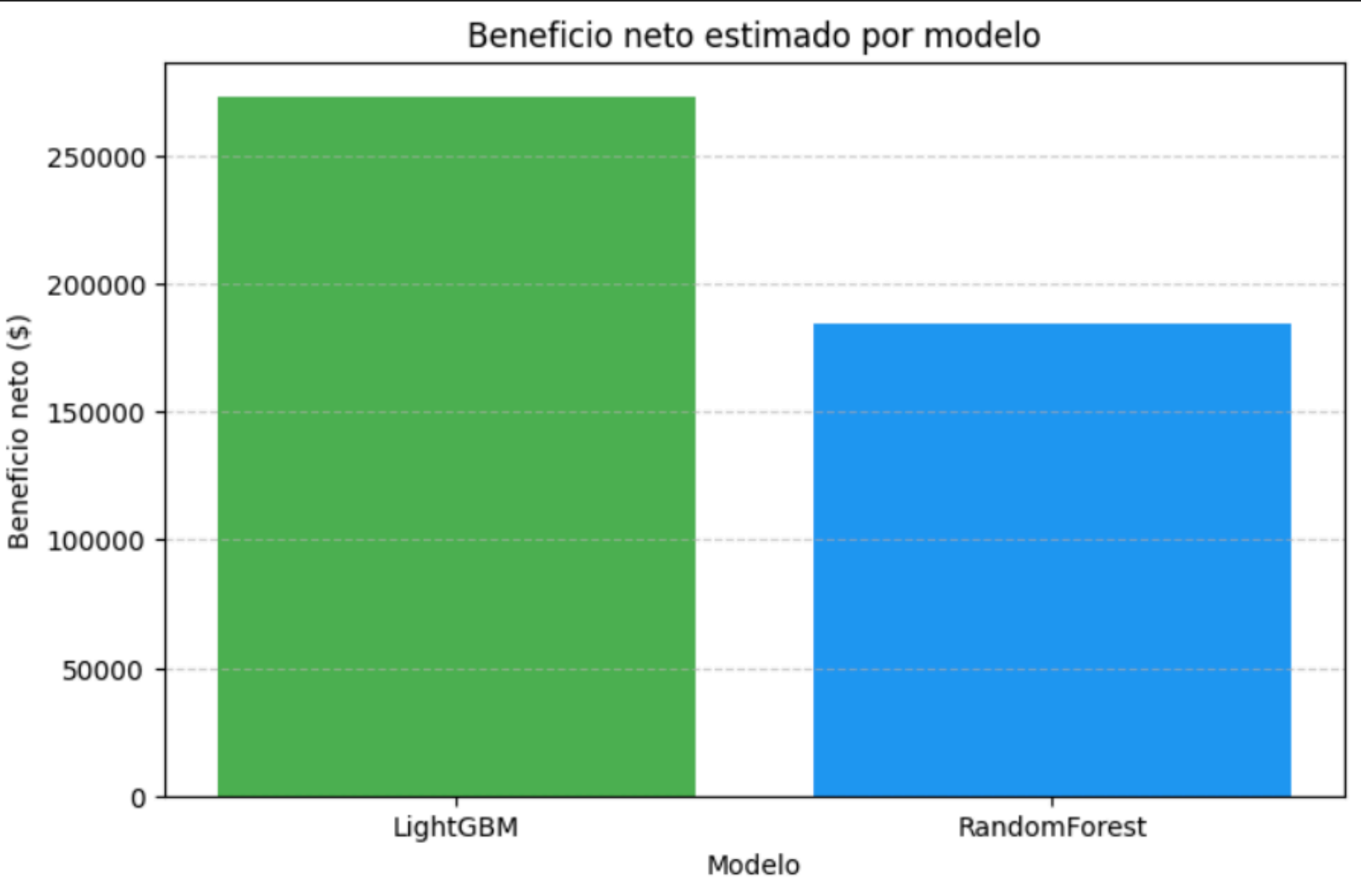


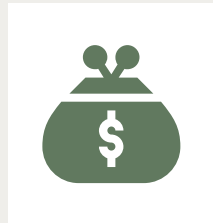
Beneficio Neto Por Modelo

📊 Comparación económico-financiera de los modelos:

	Modelo	TP (captados reales)	FP (contactos sin interés)	\
0	LightGBM	3779	17215	
1	RandomForest	2211	5159	

	Total contactados	Costo campaña (\$)	Ingresos (\$)	Beneficio neto (\$)
0	20993	104967	377880	272913
1	7370	36850	221100	184250

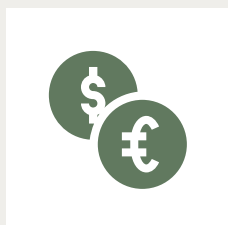


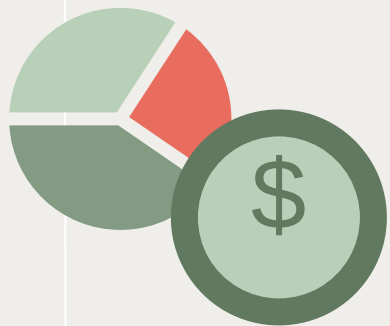


Estrategia de captación con LightGBM

Dado su altísimo recall (94%), el modelo LightGBM se alinea perfectamente con una estrategia agresiva de marketing directo donde:

- Preferimos contactar más clientes, incluso con algunos falsos positivos, asegurando así que muy pocos clientes interesados se pierdan.
- Esto permite maximizar los ingresos derivados de nuevas contrataciones, fortaleciendo el crecimiento de la cartera activa del banco.





Con base en los resultados obtenidos:

LightGBM es el modelo óptimo para implementar en la campaña de captación de préstamos y tarjetas, ya que permite capturar el 70% más de clientes interesados que Random Forest, traduciéndose en una rentabilidad neta significativamente mayor.

Adicionalmente, al tener un AUC de 0.87, garantiza un desempeño robusto ante futuros cambios en el mercado, facilitando la escalabilidad del modelo a nuevos segmentos de clientes.



Conclusión general

En un contexto altamente competitivo, donde los márgenes bancarios dependen cada vez más de la capacidad de anticipar el comportamiento de los clientes, este proyecto surge como una respuesta estratégica para transformar datos históricos en decisiones comerciales inteligentes.

Mediante el desarrollo de modelos avanzados de machine learning, se ha buscado predecir qué clientes tienen mayor probabilidad de adquirir productos financieros clave como préstamos personales o tarjetas de crédito.

Comparado con enfoques tradicionales o modelos alternativos, esta solución logra:

- Un recall de hasta el 94%, asegurando que casi ningún cliente potencial se pierda.
- Un beneficio económico neto un 47% superior, maximizando el retorno sobre cada dólar invertido en campañas.

Esto posiciona al banco a la vanguardia del uso de analytics para potenciar su crecimiento, permitiendo tomar decisiones proactivas, eficientes y completamente orientadas a resultados.

Con esta base sólida, se abren las puertas a estrategias futuras de cross-selling, personalización y fidelización que consolidarán la ventaja competitiva del banco en el mercado.

Gracias



CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)

Please keep this slide for attribution

