

Equipo: Datosos



# Datathon 2025

Reto: Análisis de cobranza domiciliada



# Objetivos principales

01.

Organizar y  
estructurar la  
información

02.

Modelo de  
cobranza

03.

Minimizar costo de  
comisiones

04.

Maximizar  
cobranza

# Arquitectura y procesamiento de datos

## Medallion Architecture



- Se **ingestaron** los datos en formato parquet en capa de bronze.
- Se **limpiaron, transformaron y se aumentaron** en la capa de silver
- Por ultimo se hicieron **agrupaciones y filtros** para llegar a business-level data en la capa gold.

### Lo que se utilizó para el modelo

Se emplea la capa gold, la cual proporciona datos depurados y listos para ser utilizados por el modelo

# Variables creadas



## Indicador de primer intento

Variable booleana que indica si el registro corresponde al primer intento de cobro realizado.



## Número de días desde último intento

Calcula la cantidad de días transcurridos entre el intento actual y el intento anterior del mismo crédito.



## Tasa de éxito previo

Porcentaje de intentos anteriores asociados al mismo crédito que tuvieron un resultado exitoso, respecto al total de intentos anteriores.



# Modelo predictivo



## Objetivo:

Identificar patrones clave y variables relevantes que no eran evidentes con análisis descriptivos tradicionales.

## Modelo elegido:

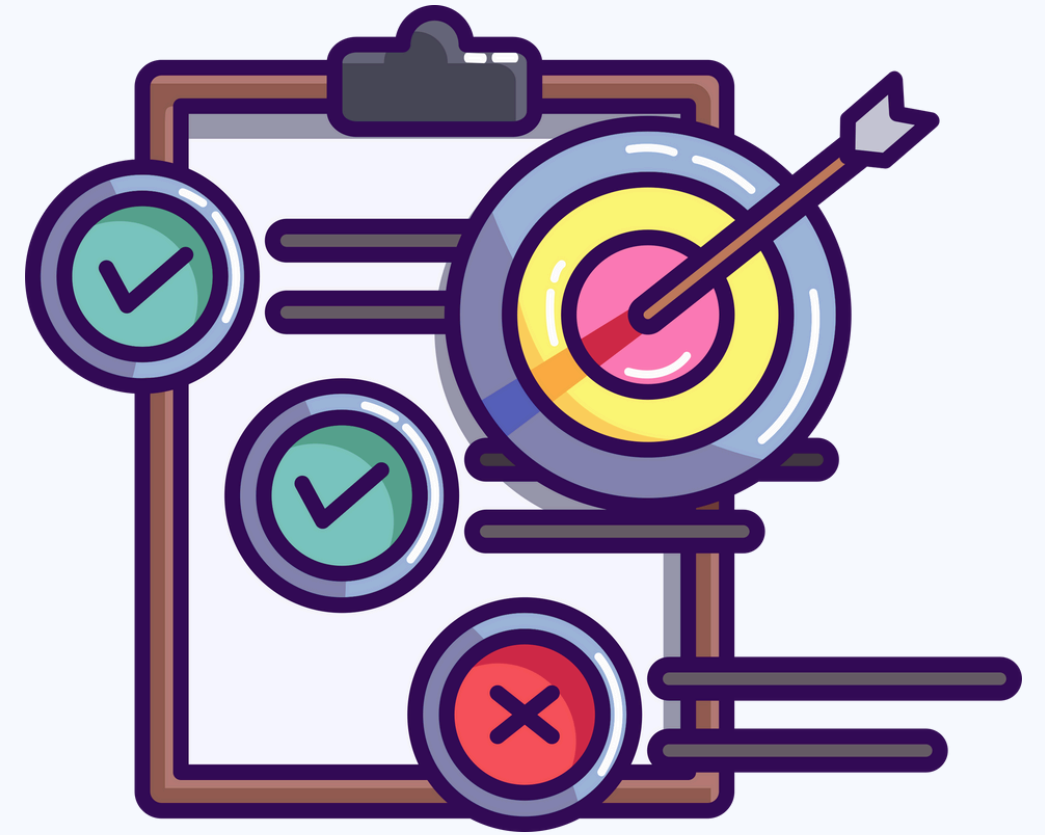
XGBoost

## Accuracy:

89.0%

# Resultados

1. Priorizar **primeros intentos** de cobro
2. Usar **emisoras con mejor historial**
3. Establecer **ventana de tiempo** óptima entre reintentos
4. El **monto no predice éxito** → enfocar esfuerzos por volumen, no por monto
5. Cliente con **historial exitoso = mejor futuro**



# Priorizar primeros intentos de cobro

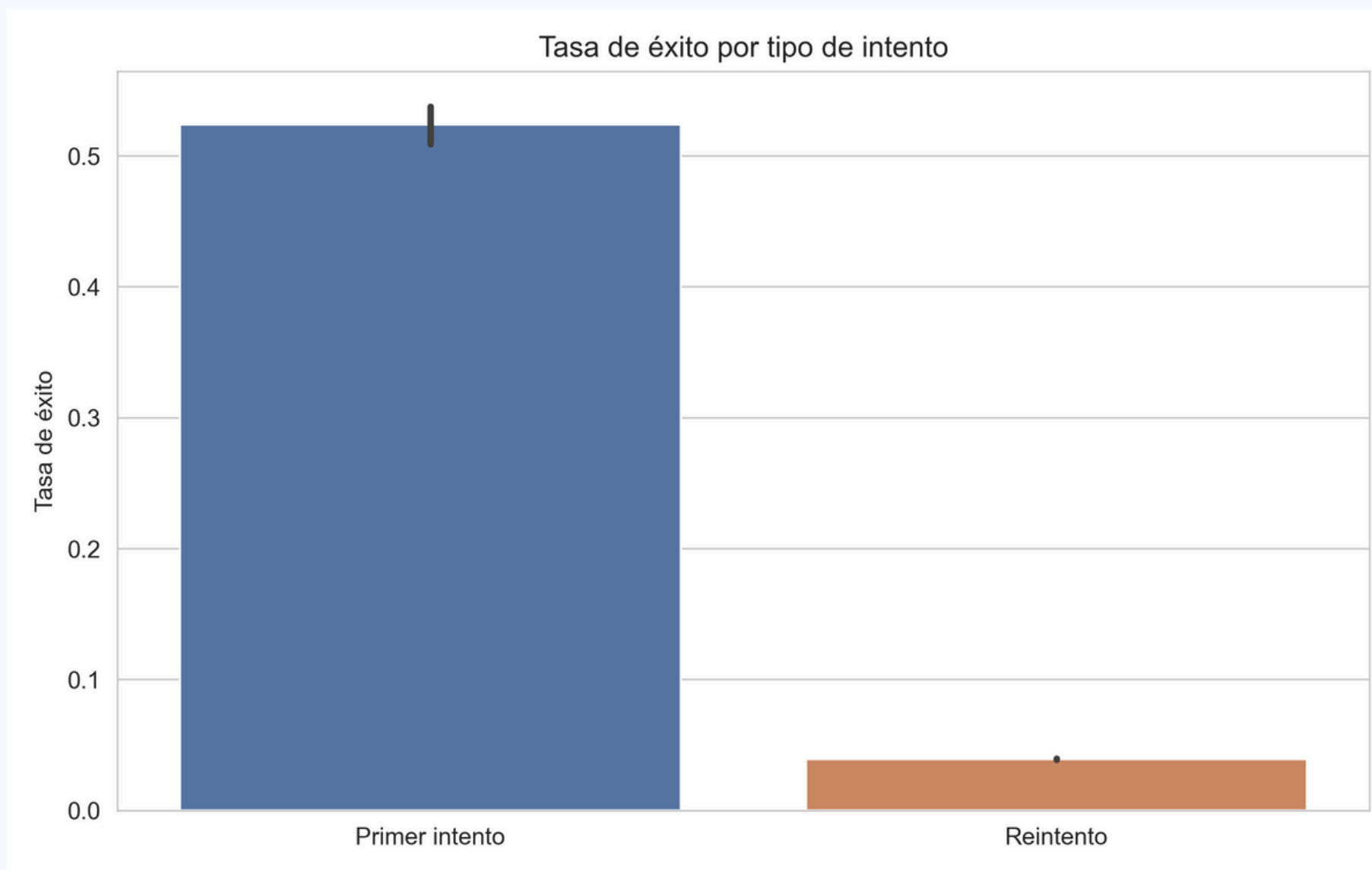
Hallazgo del modelo: ***intento\_es\_reintento*** es la variable más importante.

Regla de negocio:

*“Los primeros intentos tienen mayor probabilidad de éxito. Se deben priorizar operativamente y ejecutarse en horarios óptimos con mayor cuidado (por ejemplo, no se deben agrupar con lotes de reintentos)”.*

Acción sugerida:

- Crear una cola separada para primeros cobros.
- Aplicar mejores emisoras y estrategias ahí.



# Usar emisoras con mejor historial

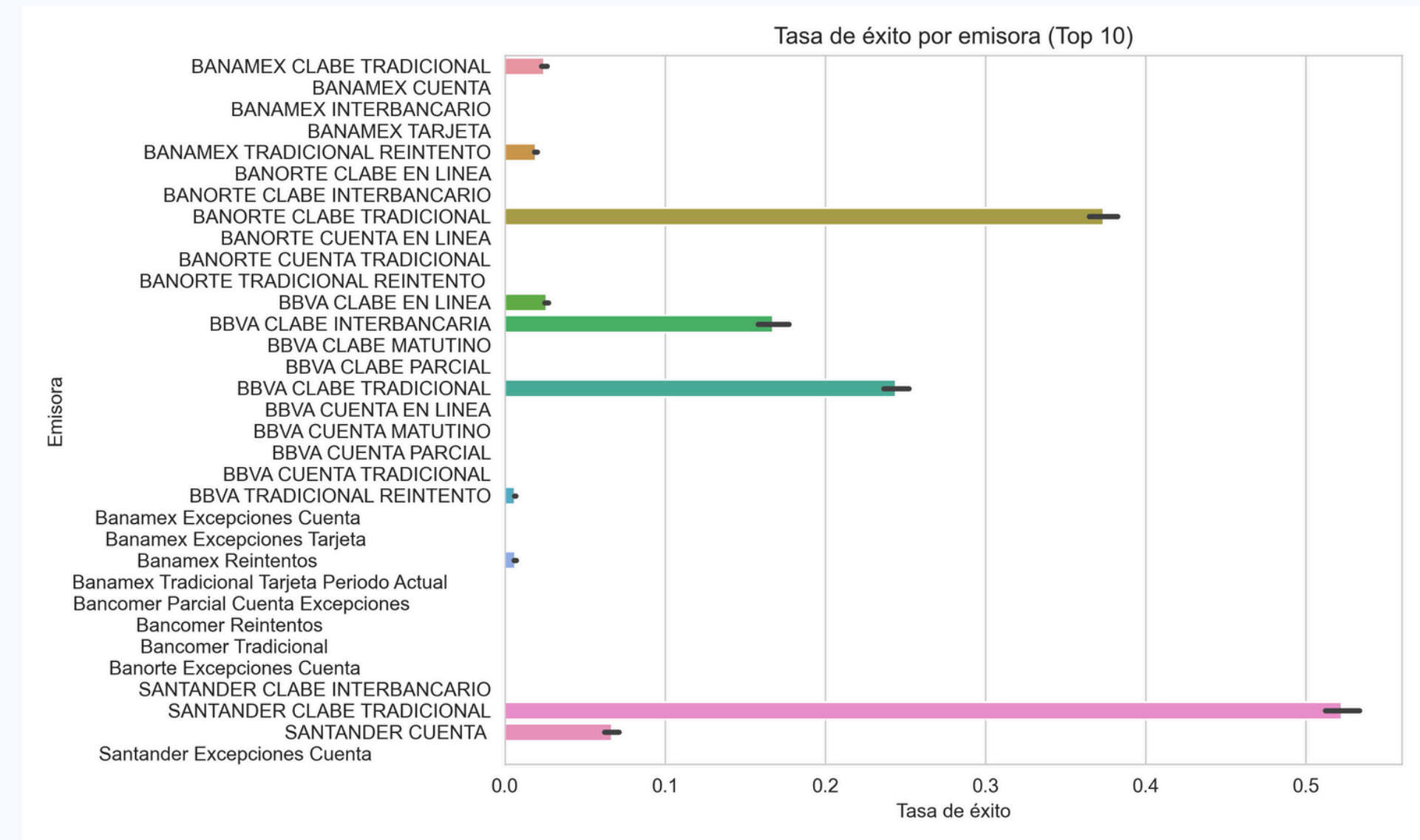
Hallazgo del modelo: **NOMBRE\_EMITORA** es la segunda variable más importante..

## Regla de negocio:

*“Algunas emisoras tienen más éxito que otras. Para créditos nuevos o cuando el primer intento falla, usar la emisora con mejor desempeño histórico del cliente”.*

## Acción sugerida:

- Generar un ranking por tasa de éxito histórica de emisoras.
- Automatizar selección prioritaria de emisora.





# Establecer ventana de tiempo óptima entre reintentas

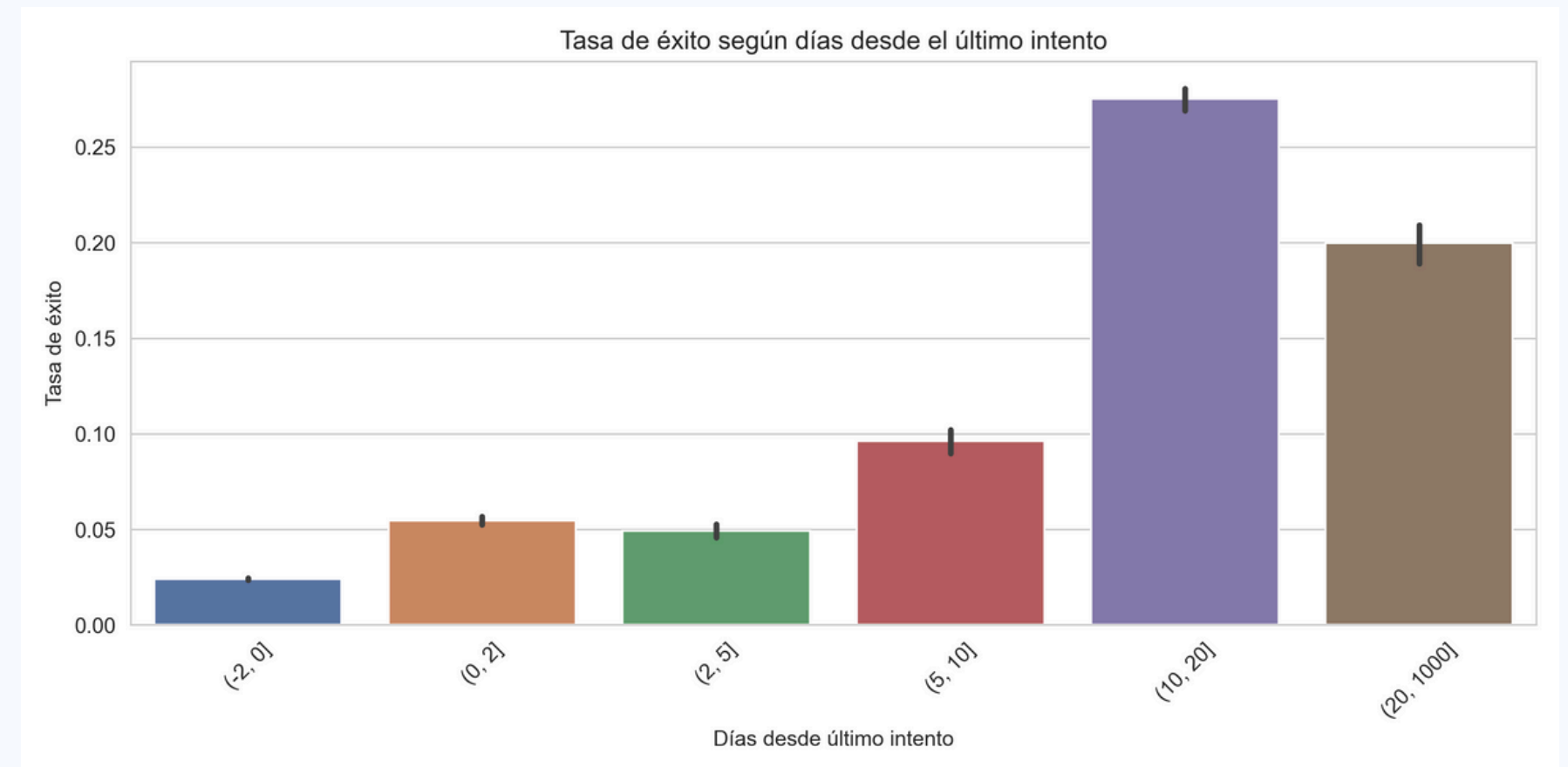
Hallazgo del modelo: ***dias\_desde\_ultimo\_intento*** es relevante.

Regla de negocio:

“El tiempo entre intentos afecta la probabilidad de éxito. Reintentar muy rápido puede ser menos efectivo que esperar cierto número de días”

Acción sugerida:

- Programar reintentos según esta lógica.



# El monto no predice éxito → enfocar esfuerzos por volumen, no por monto

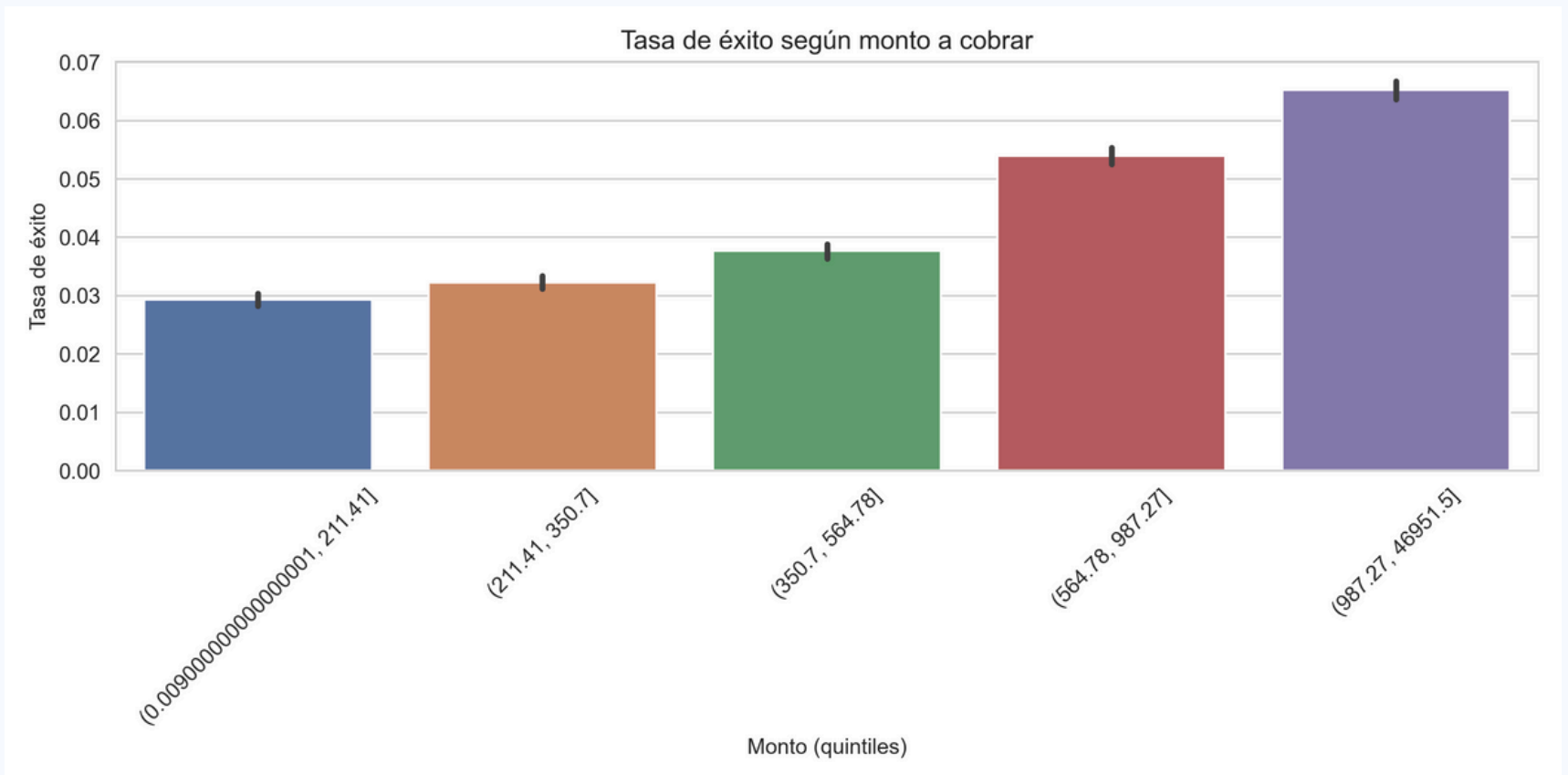
Hallazgo del modelo: **MONTOEXIGIBLE** y **MONTOCOBRAR** tienen muy baja importancia.

Regla de negocio:

*“La probabilidad de éxito no depende del monto, por lo que no conviene filtrar cobros pequeños esperando fracaso, ni sobrevalorar cobros grandes”*

Acción sugerida:

- No excluir cobros bajos del flujo operativo.
- Invertir esfuerzo similar en todos los montos si el cliente tiene buen historial.



# Cliente con historial exitoso = mejor futuro

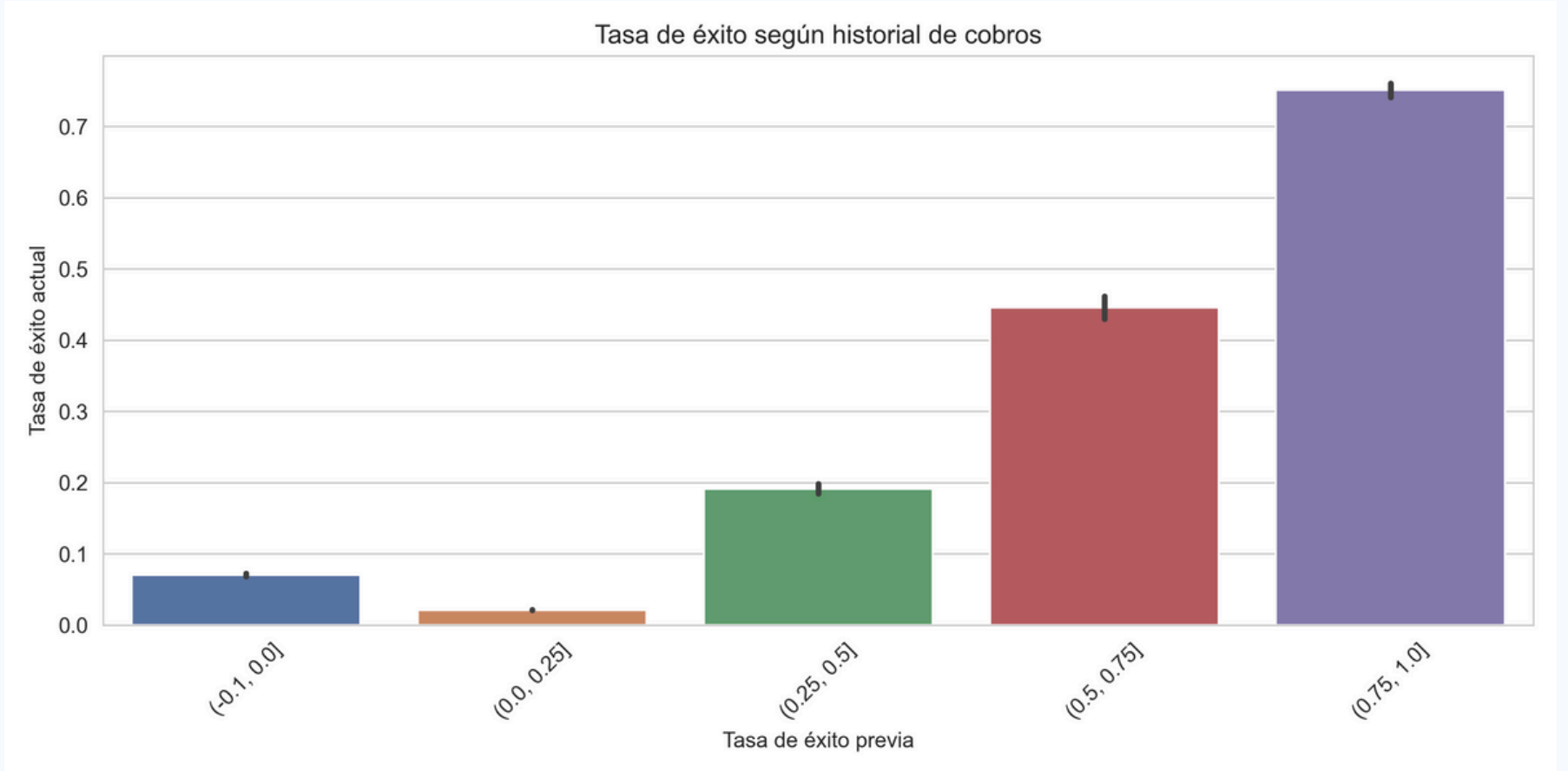
Hallazgo del modelo: ***tasa\_exito\_prev*** es relevante.

Regla de negocio:

*“Si el cliente ha pagado antes, la probabilidad de éxito se incrementa. Esto debe influir en la decisión de reintentar y cuántos intentos hacer”*

Acción sugerida:

- Configurar reintentos adicionales (o más agresivos) si el cliente tiene historial exitoso.
- Si nunca ha pagado, cortar intentos antes.



# GRACIAS

