

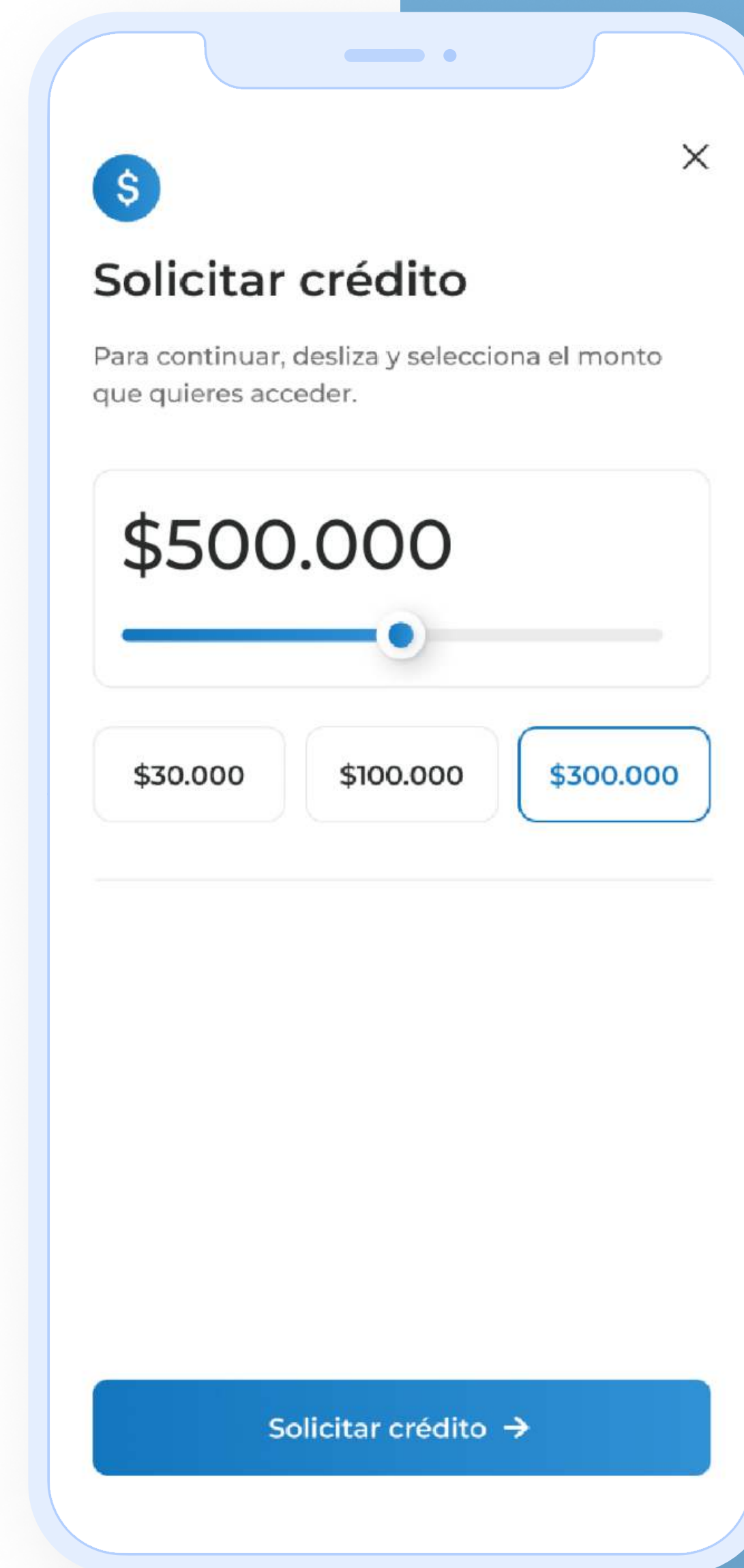
 MONET **belvo.**



CASO DE ESTUDIO MONET

Cómo Monet construyó un 'score' de crédito con datos de open finance

Monet y Belvo se asociaron para explorar el uso de datos de open finance **utilizando técnicas de machine learning**, con el objetivo de explorar los límites del 'scoring' con datos alternativos. En este caso de estudio, presentamos los resultados de esta colaboración, que muestran que los datos de open finance por sí solos tienen suficiente poder predictivo para realizar una calificación crediticia exitosa.



Contexto

Monet es una fintech especializada en préstamos de bajo monto en Colombia con la misión de desarrollar **servicios financieros inclusivos**. Los trabajadores de Latinoamérica sufren de bajo acceso al crédito y un costoso sistema financiero

tradicional. Ante esto, la solución de Monet ofrece créditos de bajo monto. La empresa ha originado más de 570,000 créditos desde junio de 2021. Desde el comienzo, su principal pilar para la aprobación de préstamos ha sido **el uso de datos de**

de open finance con sistemas basados en reglas, gracias a su colaboración con Belvo. Monet no utiliza información de burós de crédito, ya que su base de usuarios se encuentra dentro de la población sin historial crediticio o con pocos datos.

Habiendo reunido una cantidad significativa de datos a lo largo de los años, Monet decidió **explorar el uso de modelos probabilísticos** para la evaluación de nuevos préstamos.



El proyecto

El proyecto comenzó con la preparación de los datos. Los datos de open finance de Belvo se presentan en un esquema estandarizado, ya que **la plataforma se asegura de que los datos procedentes de distintas fuentes sean homogéneos**. Para expandir el espacio de variables (es decir, para ampliar los rasgos que pueden servir para caracterizar los datos) se aplicó una capa de enriquecimiento sobre la información de cuentas y transacciones. Esta capa incluyó la categorización de transacciones, la verificación de ingresos y gastos recurrentes.

Una vez que los datos estuvieron preparados, se continuó trabajando en la **creación de variables a partir de los datos de open finance** enriquecidos que luego podrían ser utilizadas para entrenamiento y predicción (es decir, 'feature engineering'). Después de varias iteraciones, se crearon más de **350 variables** que abarcan diferentes ventanas de tiempo y tipos de datos de open finance. Estas variables se pueden agrupar en seis clases:

- **Cuentas.** Utilizando información de cuentas de ahorro, cuentas de préstamos y tarjetas de crédito.
- **Saldos.** Construido a partir de cuentas y transacciones para reflejar la evolución del saldo.
- **Actividad.** Considerando patrones de actividad en fechas de valor.
- **Transaccionalidad.** Utilizando el número y tipo de transacciones históricas.
- **Caja de efectivo.** Utilizando flujos de efectivo positivos y negativos en ventanas de tiempo determinadas.
- **Categorías.** Combinando transacciones con categorización para construir perfiles de clientes.



Otra consideración importante fue la **definición de buenos y malos pagadores** que se utilizarían en el entrenamiento del algoritmo. Dado el modelo de negocio de Monet, se consideró un incumplimiento de pago temprano como un horizonte de tiempo de desempeño adecuado.

Para la definición de incumplimiento, se eligió el nivel de días de atraso teniendo en cuenta las curvas de recuperación de cohortes anteriores.

Con el conjunto de datos final en mano, el proyecto pasó a la fase de **aprendizaje supervisado**.

El clasificador de aprendizaje automático elegido fue XGBoost y el conjunto de datos se dividió en una muestra de entrenamiento (80%) para la selección de hiperparámetros y características, y una muestra de prueba (20%) para la **validación del rendimiento**.

Los hiperparámetros se ajustaron utilizando una búsqueda aleatoria configurada para clasificación desbalanceada y las características se seleccionaron utilizando valores SHAP.

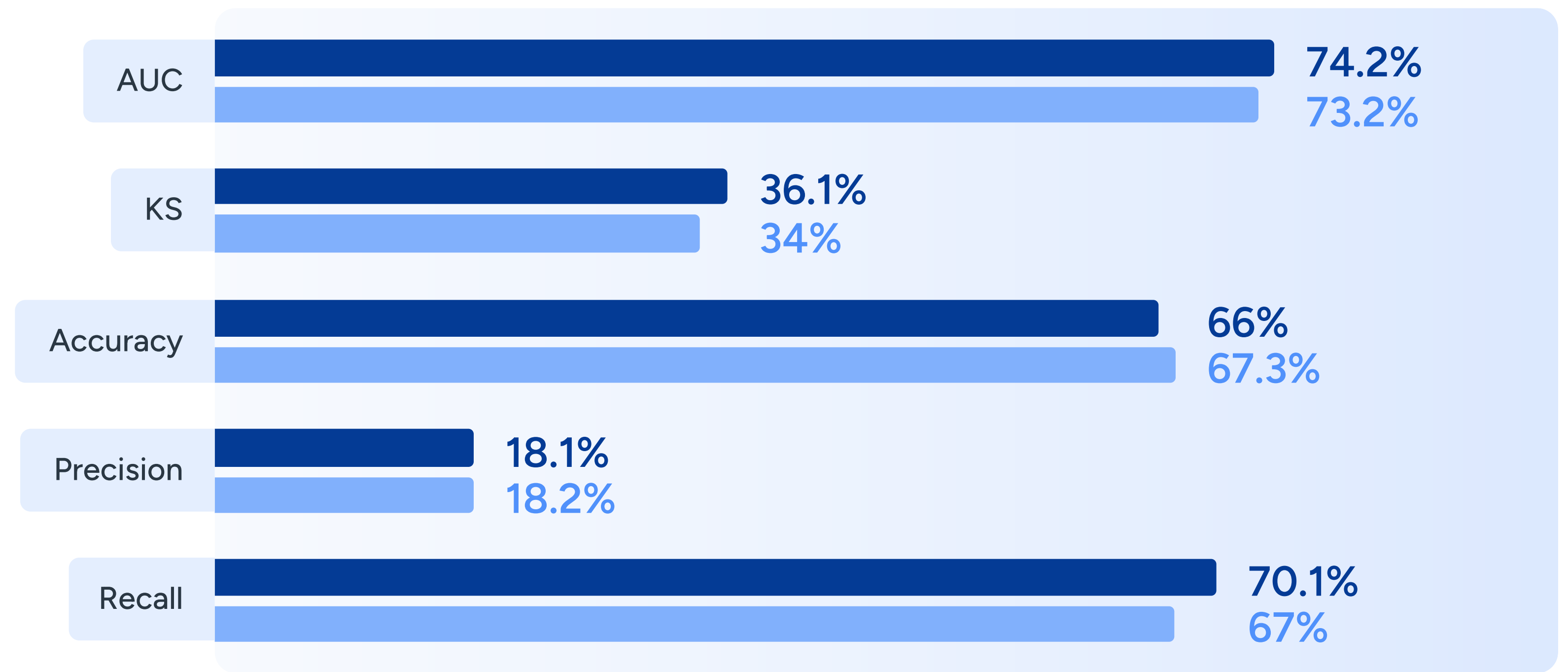
Los resultados

Desde el punto de vista de la clasificación, el rendimiento del modelo de puntuación se consideró exitoso con un AUC del 73% y un KS del 34% en la muestra de prueba.

AUC ('Area under the curve') y KS (Kolmogorov-Smirnov test) son medidas de calidad de scoring para entender hasta qué punto el modelo es capaz de distinguir entre clases.

El gráfico muestra las métricas de clasificación comunes tanto para la muestra de entrenamiento como para la de prueba:

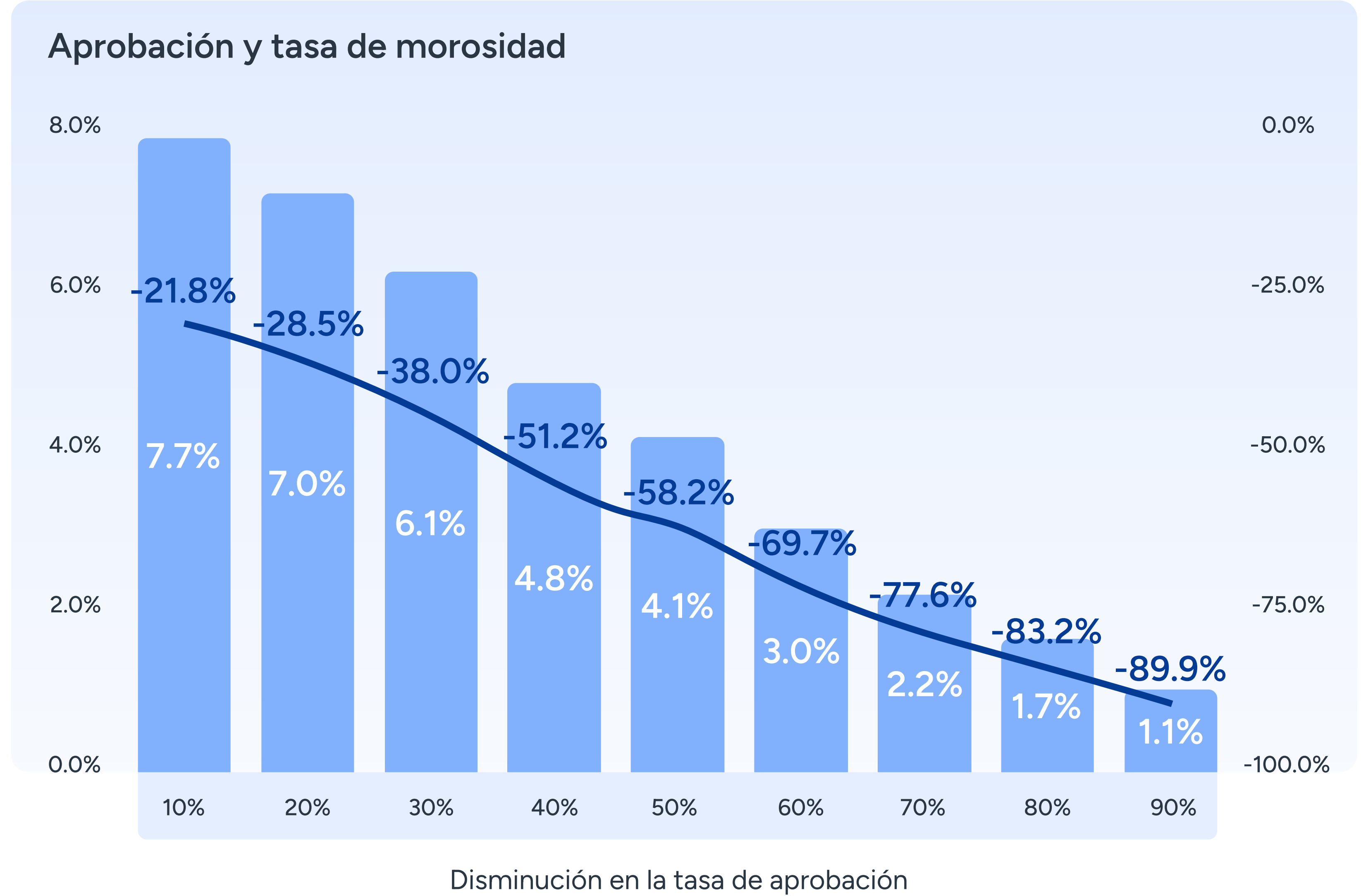
■ Entrenamiento ■ Prueba



Desde el punto de vista económico, también se consideró exitoso el rendimiento del modelo de calificación.

La empresa ahorrará más del 20% en morosidad utilizando modelos probabilísticos.

El gráfico a continuación muestra la reducción en las tasas de morosidad que se podrían lograr al cambiar los umbrales de aprobación:



■ Tasa de morosidad

— Cambio en la tasa de morosidad

belvo.

¿Estás buscando
mejorar tus modelos de
'scoring' de crédito?

Contáctanos: hello@belvo.com