Deriva

Una alerta de deriva es el inicio de la investigación.

Problema operativo

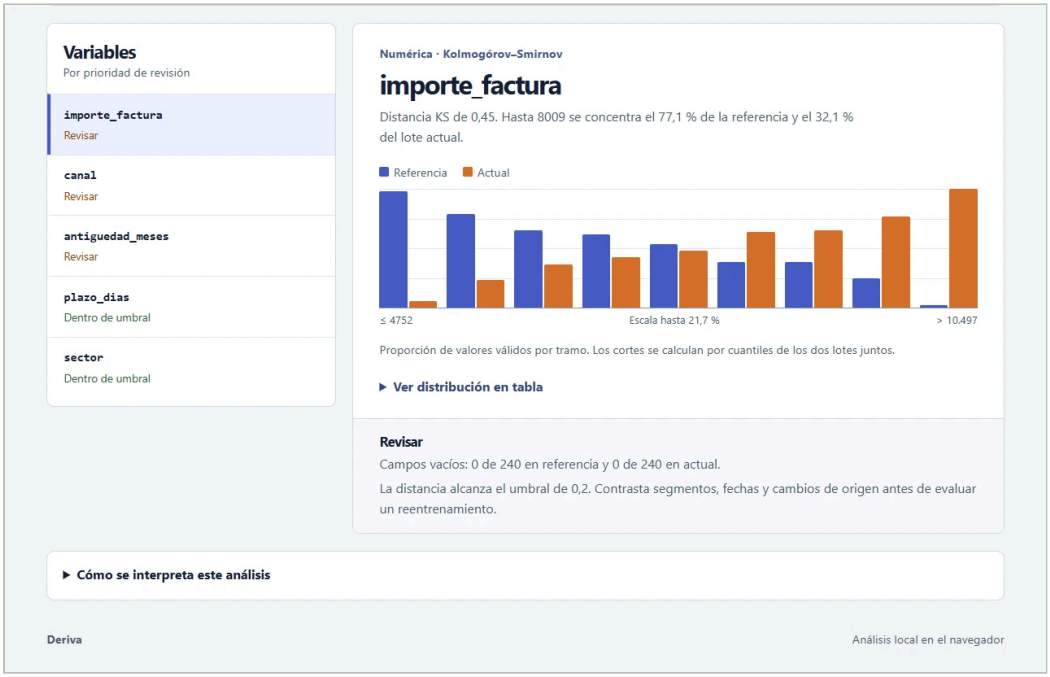
Suben los importes de las solicitudes de financiación. El equipo de riesgo necesita localizar en qué tramo aparece el cambio y comprobar si también varían los datos ausentes antes de interpretar el comportamiento del modelo de crédito.

Qué desarrollé

Implementé comparaciones Kolmogórov–Smirnov y Jensen–Shannon por tipo de variable, con un historial de informes en Python y SQLite. La recalibración de la probabilidad de impago (PD) usa las etiquetas disponibles antes de la cohorte evaluada.

Resultados y alcance

En el escenario sintético, un ajuste de junio reduce el error Brier de 0,1556 a 0,1503 en agosto. En septiembre, el mismo ajuste lo eleva de 0,1072 a 0,1132. En los datos de campañas bancarias de UCI, la categoría de contacto móvil figura en 0 de los primeros 1.000 registros y en 817 de los últimos 1.000.



La distancia KS de 0,45 señala dónde más se separan los importes. En ese punto, la proporción acumulada de solicitudes difiere en 45 puntos porcentuales.

Tecnologías Python · SQLite · Node.js · Estadística
Oferta investigada Novicap · Data & Machine Learning Engineer
Comprobar si un ajuste de PD mejora en una cohorte posterior Proyecto · rcr-ia.eu/trabajo/deriva/
Ricardo Crespillo · LinkedIn