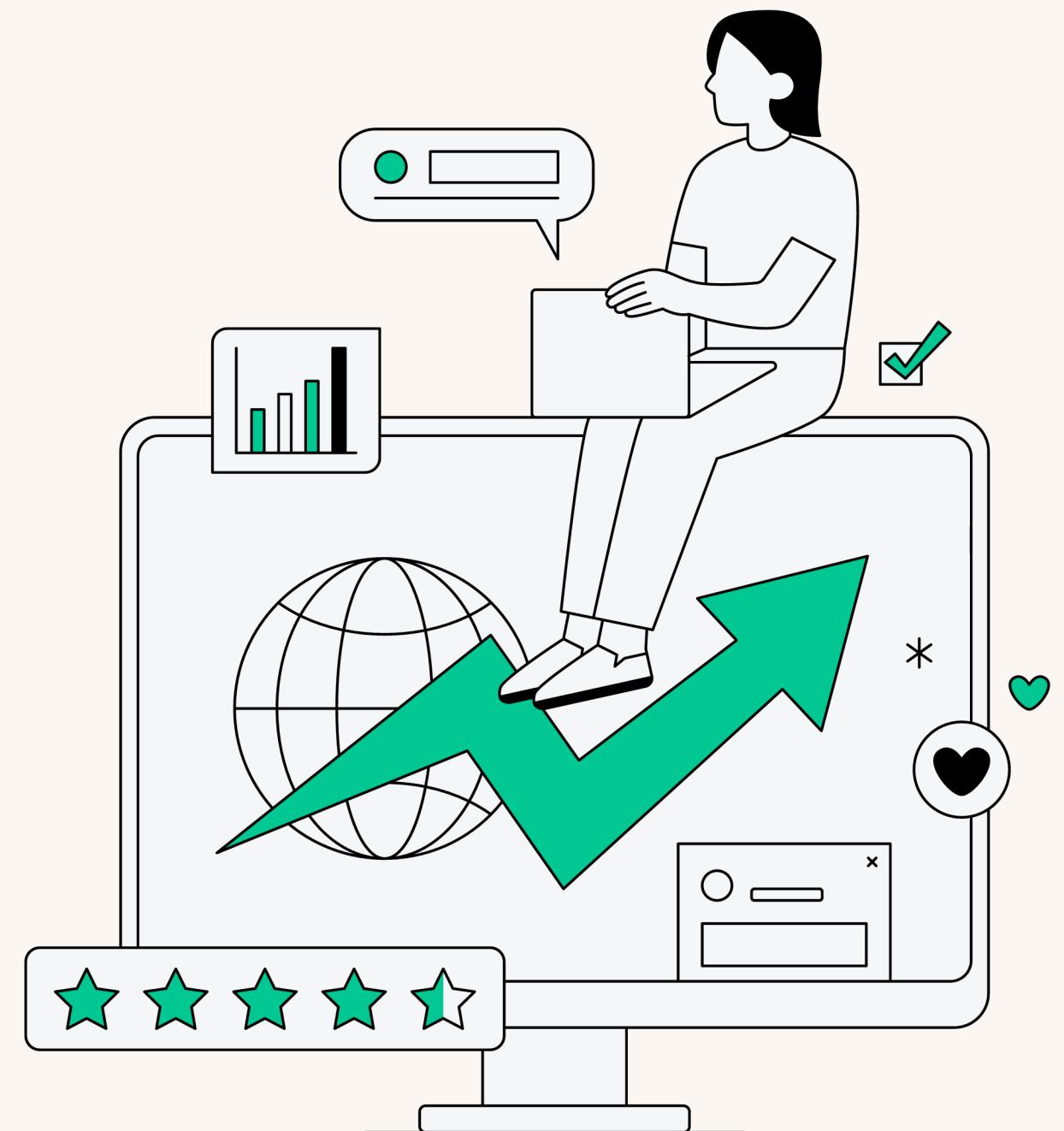


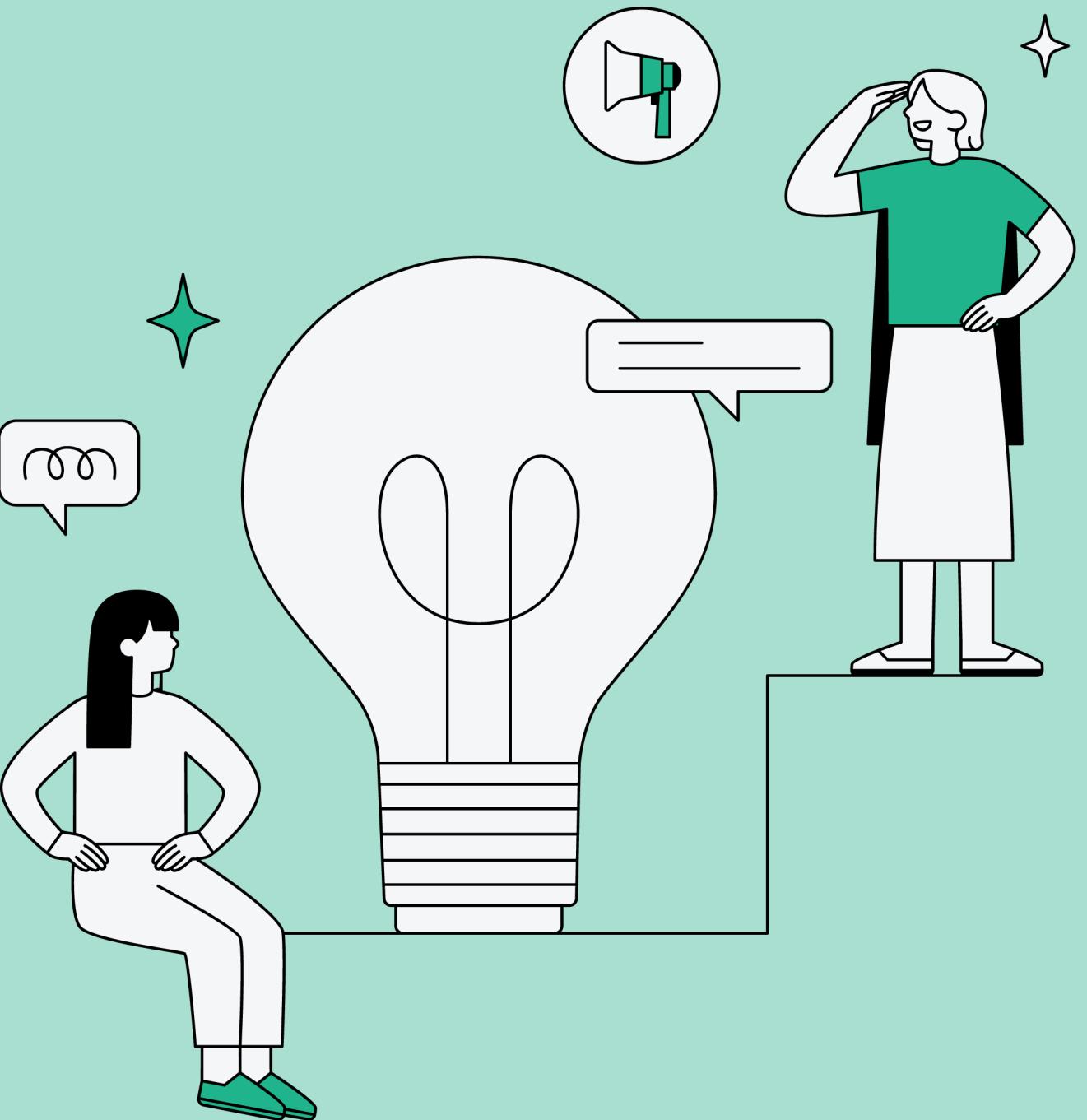
Predicción de disputas de clientes con ML

Presentado por Anthony Quiliche



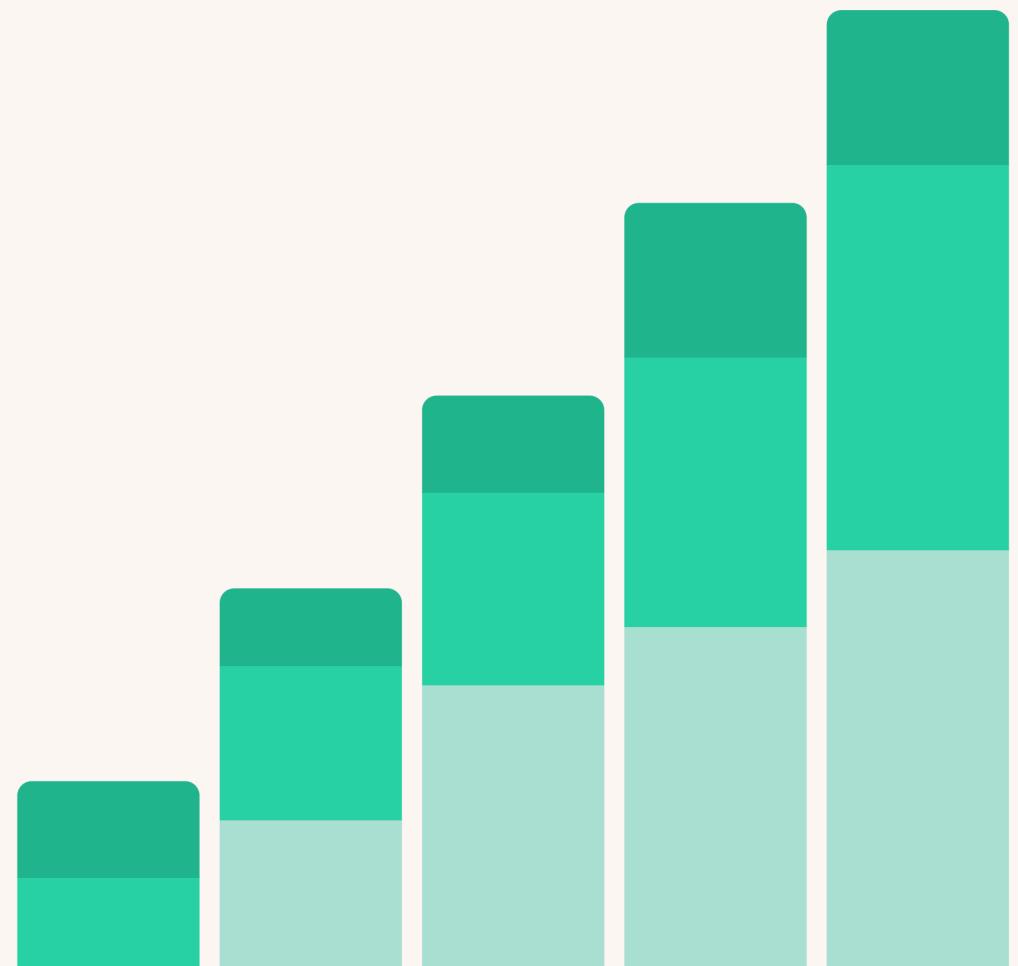
Contexto del Problema

- Problema: “Actualmente, las disputas generan costos adicionales y retrasos en la resolución”.
- Solución: “Un modelo predictivo que identifica quejas con alta probabilidad de disputa”.



Flujo del proyecto

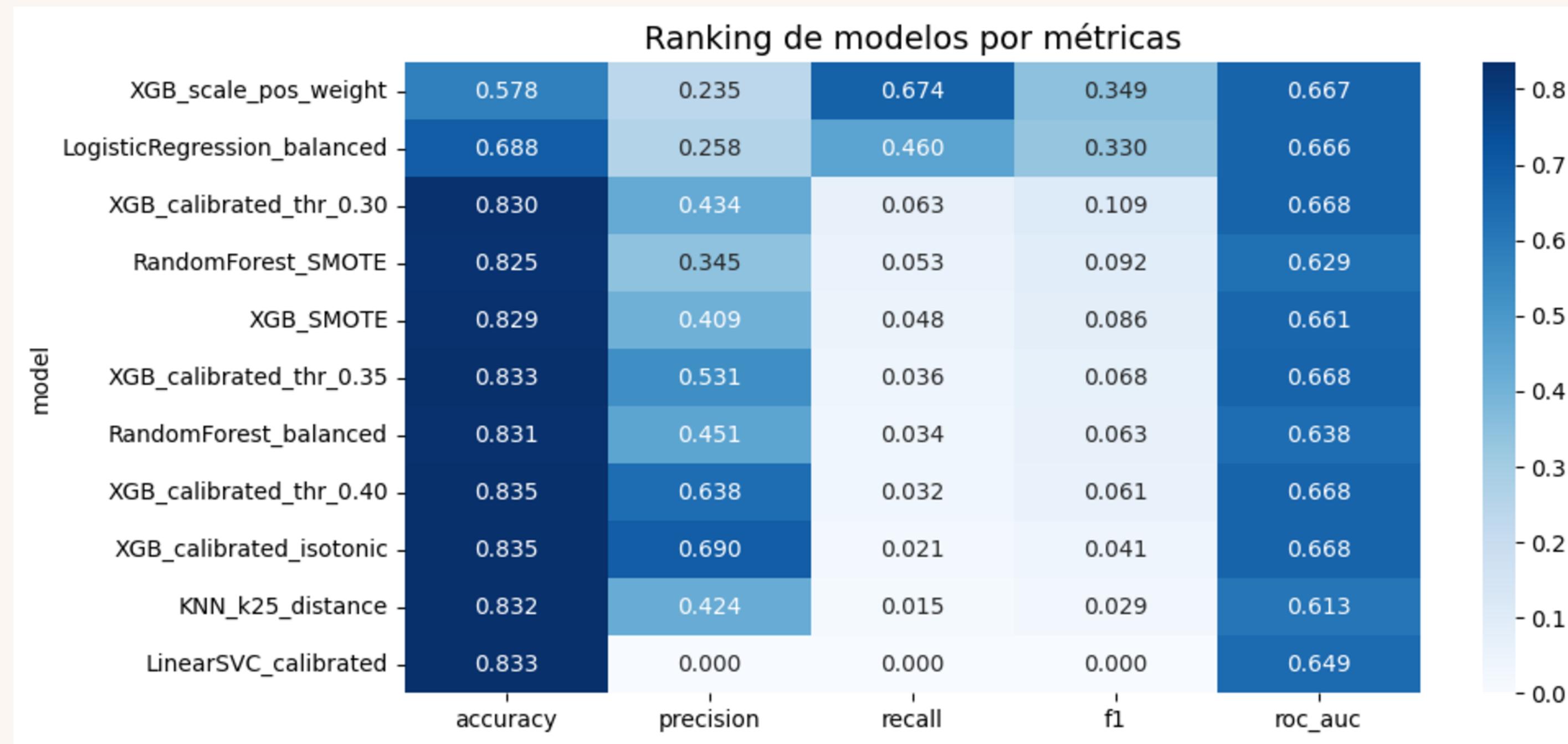
- 1. Recolección y limpieza de datos (detección de nulos, outliers).
- 2. Análisis exploratorio (EDA) para identificar patrones.
- 3. Ingeniería de características y encoding de variables categóricas.
- 4. Entrenamiento con varios modelos (Logistic, Random Forest, XGBoost).
- 5. Evaluación con métricas F1, AUC, Recall y Precision.

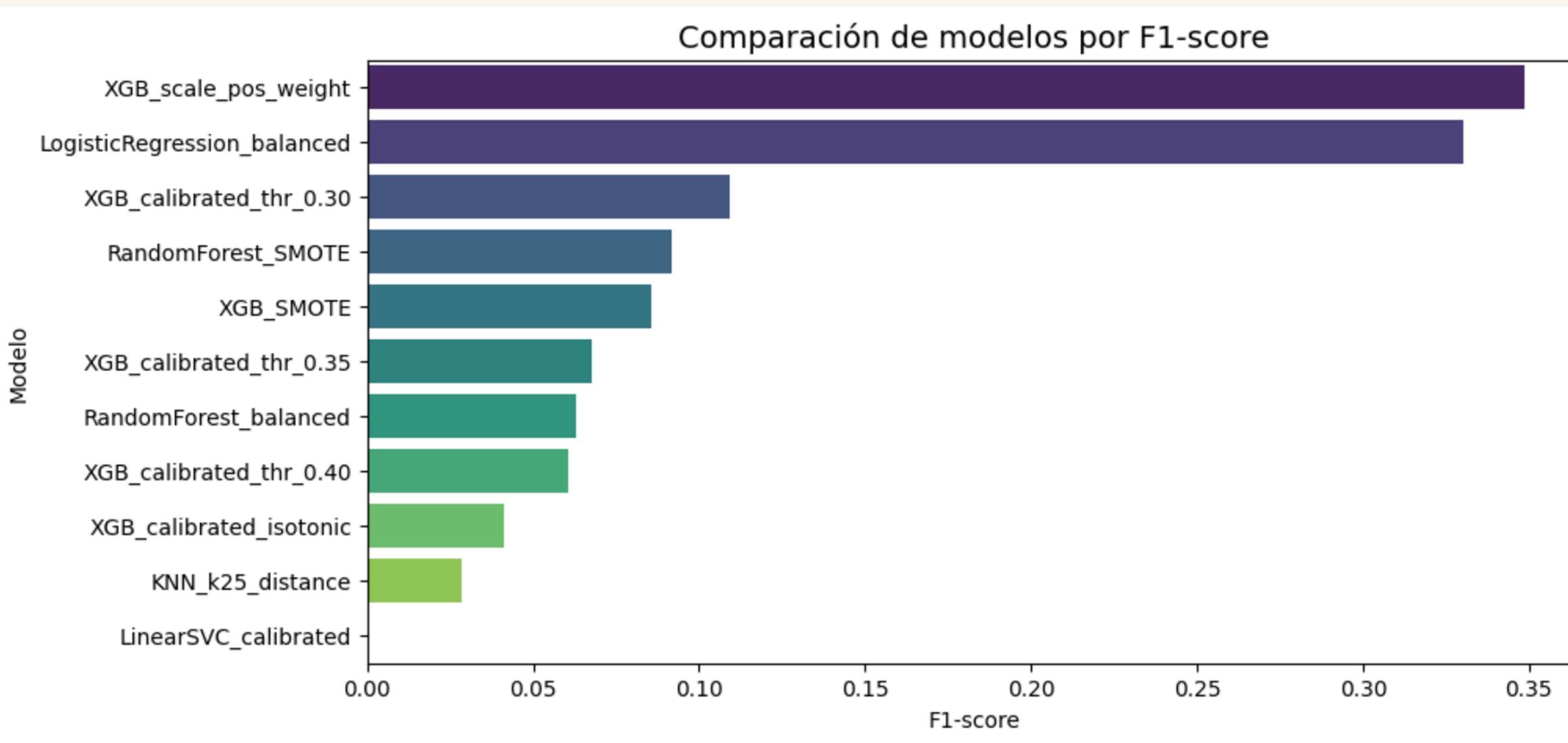


encoding de variables categóricas.

	Product	Sub_product	Issue	Sub_issue	State	ZIP_code	Date_received	Date_sent_to_company	Company	Company_response	Timely_response	Consumer_disputed
0	4	Desconocido	Communication tactics	Frequent or repeated calls	51	76119.0	2015-03-19	2015-03-19	1069	4	1	0
1	4	Medical	Cont'd attempts collect debt not owed	Debt is not mine	51	77479.0	2015-03-19	2015-03-19	50	1	1	0
2	6	FHA mortgage	Application, originator, mortgage broker	Desconocido	24	2127.0	2015-03-19	2015-03-19	1118	1	1	1
3	2	Desconocido	Other	Desconocido	7	92592.0	2015-03-19	2015-03-19	967	4	1	0
4	4	Non-federal student loan	Cont'd attempts collect debt not owed	Debt resulted from identity theft	12	43068.0	2015-03-19	2015-03-19	1401	4	1	0
5	4	Payday loan	Communication tactics	Called after sent written cease of comm	41	43068.0	2015-03-19	2015-03-19	15	4	1	0
6	1	Installment loan	Managing the loan or lease	Desconocido	41	44241.0	2015-03-18	2015-03-18	291	1	1	0
7	4	Payday loan	Communication tactics	Called after sent written cease of comm	7	95124.0	2015-03-18	2015-03-18	15	1	1	0
8	4	Other (phone, health club, etc.)	Cont'd attempts collect debt not owed	Debt was paid	37	7067.0	2015-03-18	2015-03-18	409	3		

Entrenamiento de modelos





Modelo elegido → XGBoost

- Excelente manejo de variables categóricas y numéricas.
- Mejor rendimiento en recall y F1 frente a otros modelos.
 - Robusto ante datos desbalanceados.
 - Permite interpretabilidad mediante SHAP values.
 - Balance ideal entre precisión y explicabilidad.

App Streamlit

La aplicación permite realizar predicciones tanto individuales como masivas sobre el riesgo de disputa en quejas de clientes, mostrando no solo el resultado numérico, sino también explicaciones visuales con métricas e interpretabilidad. Publicado en la plataforma Render.

<https://project-machinelearning-client.onrender.com>



Conclusiones

El modelo XGBoost permite anticipar disputas de clientes con alta precisión.

El dashboard facilita el uso por áreas técnicas y de negocio.

Proyecto escalable para integrarse en flujos reales de atención.

Aporta valor práctico para el negocio.



Gracias

Presentado por Anthony Quiliche

