

Reporte de Modelos – Proyecto Integrador FinanceGuard

Cierre del Proyecto: Reducción de Churn en Banca Digital

Documento técnico final

Avance 4 – Integración de hallazgos y recomendaciones estratégicas

Equipo de Retención – FinanceGuard

1. Resumen Ejecutivo

FinanceGuard enfrenta una **tasa de churn del 20,4%** anual. A lo largo de tres avances se desarrollaron modelos supervisados (Regresión Logística y Gradient Boosting) y análisis no supervisado (clustering y reducción de dimensionalidad) para entender y predecir el abandono de clientes.

Este reporte consolida los resultados de cada avance, compara el rendimiento de los modelos, sintetiza los aprendizajes y presenta recomendaciones estratégicas para el equipo de retención.

Hallazgos principales:

- **Mejor modelo supervisado:** Stacking (XGBoost + LightGBM) con **ROC-AUC 0,872** y **F1 0,592**.
 - **Ganancia vs baseline:** El Gradient Boosting (XGBoost optimizado) supera a la Regresión Logística en discriminación (ROC-AUC ~0,87 vs ~0,84) y en capacidad de capturar relaciones no lineales.
 - **Segmentación:** K-Means con K=3 identifica un cluster de alto riesgo (Alemania, **32,4% churn**) frente a ~16% en otros segmentos.
 - **Variables clave:** Edad, NumOfProducts, Balance e IsActiveMember son los predictores más importantes tanto en modelos supervisados como en perfiles de cluster.
-

2. Síntesis por Avance

2.1 Avance 1 – Regresión Logística

Objetivo: Establecer un modelo baseline interpretable para predecir churn.

Performance del modelo baseline

Métrica	Valor (aprox.)	Interpretación
Accuracy	~80–82%	Correcto en 8 de cada 10 predicciones
ROC-AUC	~0,84	Buena capacidad de discriminación
Precision	~0,65–0,70	De los alertados como churn, ~65–70% realmente se van
Recall	~0,35–0,45	Detecta ~35–45% de los clientes que se van
F1-Score	~0,45–0,55	Balance entre precisión y recall

- Dataset: **10.000 clientes**, 20,4% churn (desbalance 80/20).
- Variables: CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrard, IsActiveMember, EstimatedSalary (con encoding de Geography y Gender).

Interpretabilidad y coeficientes más importantes

- **Age (positivo)**: A mayor edad, mayor probabilidad de churn (odds ratio > 1).
- **Geography_Germany (positivo)**: Clientes en Alemania tienen mayor riesgo que la referencia (Francia).
- **Balance (positivo)**: Saldo alto asociado a mayor churn (posible efecto “preparación para salida”).
- **IsActiveMember (negativo)**: Miembros activos tienen menor probabilidad de churn.
- **NumOfProducts**: Efecto no lineal; 1 o 4 productos asociados a mayor churn que 2–3.

Fortalezas

- Fácil de explicar al negocio (coeficientes y odds ratios).
- Rápido de entrenar y desplegar.
- Proporciona probabilidades calibradas para priorizar acciones.
- Sirve como referencia para modelos más complejos.

Limitaciones

- Relaciones lineales en el espacio log-odds; puede subestimar interacciones.
- Recall moderado con umbral 0,5: se pierde una parte de los churners.
- No captura bien patrones no lineales que sí aprovechan los árboles y boosting.

2.2 Avance 2 – Gradient Boosting y Optimización

Objetivo: Mejorar la predicción de churn con modelos de ensemble y optimización de hiperparámetros.

Mejor modelo identificado

- **Stacking** (XGBoost + LightGBM con meta-learner Logístico): mejor **ROC-AUC (~0,872)** en comparación con los demás.
- **XGBoost con Grid Search** es la base principal: ROC-AUC ~0,871, Accuracy ~86,7%, F1 ~0,59.

Parámetros óptimos (XGBoost, ejemplo representativo)

- max_depth: 3
- learning_rate: 0,1
- n_estimators: 100
- subsample: 0,8
- colsample_bytree: 1,0
- reg_alpha: 0, reg_lambda: 1

Feature importance del mejor modelo (XGBoost / Stacking)

Orden típico de importancia (gain):

1. **Age** – Edad como principal predictor de riesgo.
2. **NumOfProducts** – Número de productos (1 o 4 más riesgosos).
3. **Balance** – Saldo en cuenta.
4. **IsActiveMember** – Actividad del cliente.
5. **Geography** (Germany/Spain) – Mercado.
6. **CreditScore, Tenure, EstimatedSalary, HasCrCard, Gender** – Contribución menor pero no nula.

Ganancia en performance vs modelo baseline

Métrica	Regresión Logística (Avance 1)	XGBoost / Stacking (Avance 2)	Mejora
ROC-AUC	~0,84	~0,87	+~3–4%
F1-Score	~0,45–0,52	~0,59	+~15–20%
Recall	~0,35–0,45	~0,47–0,50	Aumento en detección de churners

- El boosting captura **interacciones y no linealidades** (ej. Age × Balance, Geography × NumOfProducts) que la regresión logística no modela explícitamente.
- **Validación cruzada estratificada (5-fold)** con métrica ROC-AUC asegura evaluación estable en datos desbalanceados.

2.3 Avance 3 – Aprendizaje No Supervisado

Objetivo: Segmentar clientes sin usar la etiqueta de churn para descubrir perfiles y cruzar después con churn.

Segmentos de clientes identificados (K-Means, K=3)

- **Cluster 0 (Alemania, alto balance):** ~2.509 clientes; Balance alto (~119.730 USD); 100% Germany; **Churn 32,4%** (cluster de mayor riesgo).
- **Cluster 1 (Francia):** ~5.014 clientes; Balance medio (~62.093 USD); **Churn 16,2%**.
- **Cluster 2 (España):** ~2.477 clientes; Balance medio (~61.818 USD); **Churn 16,7%**.

Insights de negocio por cluster

- **Cluster 0:** Prioridad máxima para retención; concentrar ofertas de retención, revisión de experiencia en Alemania y posible ajuste de producto/precio.
- **Clusters 1 y 2:** Churn por debajo de la media; útiles para estrategias de crecimiento y cross-selling sin descuidar retención básica.
- **DBSCAN** en este dataset generó muchos outliers (datos homogéneos); K-Means resultó más útil para segmentación accionable.

Features derivadas del clustering

- **cluster_kmeans:** Etiqueta de cluster (0, 1, 2) para uso en modelos supervisados o reglas de negocio.
- **cluster_risk_rank:** Orden del cluster por tasa de churn (0 = mayor riesgo, 2 = menor riesgo), útil como variable o para priorización.

Lecciones del no supervisado

- La segmentación **no sustituye** al modelo de churn; lo **complementa** (perfiles vs probabilidad individual).
- Silhouette bajo (~0,13) indica clusters con solapamiento; aun así, la tasa de churn por cluster es **estadísticamente significativa** (chi-cuadrado).
- PCA y t-SNE mostraron que Alemania (Cluster 0) se separa algo más en el espacio de características, coherente con mayor churn y perfil de balance.

3. Lecciones Aprendidas

3.1 ¿Cuándo usar modelos supervisados vs no supervisados?

Criterio	Supervisado (Regresión Logística, XGBoost, Stacking)	No supervisado (K-Means, PCA)
Objetivo	Predecir quién hace churn y probabilidad	Entender grupos y perfiles
Uso de la etiqueta	Sí (Exited)	No (solo para análisis posterior)
Salida principal	Probabilidad de churn por cliente	Segmentos y perfiles promedios
Acción típica	Alertas y campañas por riesgo (umbral de probabilidad)	Estrategias por segmento (ej. Alemania)
Interpretabilidad	Coefficientes (logística) o importance (boosting)	Centroides y tasas por cluster

Recomendación: Usar **supervisado para scoring y priorización diaria** (ej. lista de clientes con $P(\text{churn}) > 0,5$) y **no supervisado para diseño de estrategias por segmento** (ej. producto o mensaje diferenciado por cluster/país).

3.2 Consideraciones para futuros proyectos de churn

- **Desbalance:** Mantener métricas como ROC-AUC, F1 y Recall; no depender solo de Accuracy. Considerar muestreo o pesos por clase si se requiere más recall.
 - **Umbral de decisión:** El umbral 0,5 no es sagrado; optimizar según **coste de FN** (no detectar churn) vs **coste de FP** (falsas alarmas), tal como se explora en el Extra Credit (optimización de threshold y matriz de confusión con costos).
 - **Estabilidad:** Validación cruzada estratificada y, si es posible, validación temporal para evitar overfitting y drift.
 - **Interpretabilidad:** En entornos regulados o con equipos de negocio, combinar modelos interpretables (regresión logística, árboles pequeños) con modelos de mejor rendimiento (XGBoost) mediante reportes de importancia y perfiles.
 - **Integración con clustering:** Incluir **cluster o risk_rank** como input en un modelo supervisado (o en reglas de negocio) puede mejorar la segmentación de acciones (ej. umbral más agresivo en Cluster 0).
-

4. Recomendaciones Estratégicas

1. **Desplegar el mejor modelo supervisado (Stacking o XGBoost)** para scoring semanal o diario de clientes, con lista priorizada por probabilidad de churn.
 2. **Optimizar el umbral de clasificación** según costes reales de FN y FP (ver notebook 4_Extra_credit.ipynb).
 3. **Priorizar Alemania (Cluster 0):** Acciones específicas de retención, investigación cualitativa y posible ajuste de producto/canal en ese mercado.
 4. **Variables accionables:** Fomentar uso de productos (NumOfProducts 2–3), engagement (IsActiveMember) y atención a clientes con balance alto y edad avanzada.
 5. **Monitoreo:** Seguimiento de ROC-AUC, F1 y distribución de scores en producción; alertas si la distribución de churn o de features cambia (drift).
 6. **Cierre del ciclo:** Usar resultados de campañas de retención para reentrenar modelos y recalibrar umbrales de forma periódica.
-

5. Referencia a Materiales del Proyecto

- **Avance 1:**
clase_03_regresion_logistica/notebooks/actividad_clase_03_regresion_logistica.ipynb
 - **Avance 2:**
clase_06_optimizacion_modelos/notebooks/2_GradientBoosting_Optimizacion_RESUMEN.ipynb
 - **Avance 3:**
clase_08_aprendizaje_no_supervisado_ii/notebooks/3_AprendizajeNoSupervisado.ipynb
 - **Extra Credit (umbral y costes):**
clase_09_series_temporales/notebooks/4_Extra_credit.ipynb
-