

Informe técnico: Evaluación del modelo de riesgo cardiovascular

EDA y limpieza

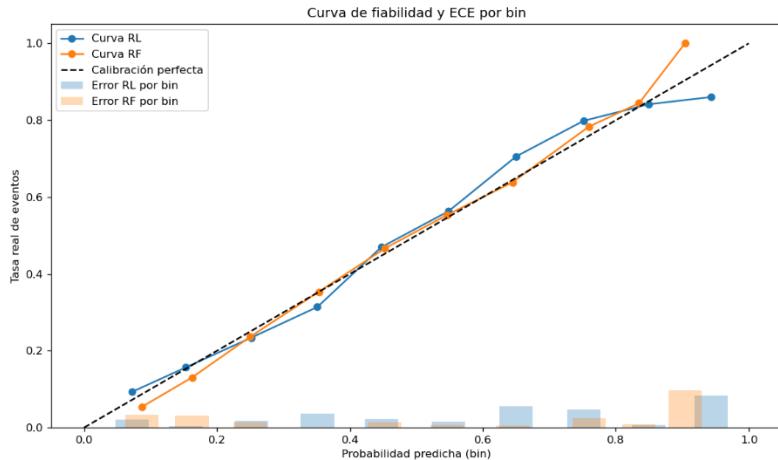
Se aplicaron reglas clínicas y técnicas para garantizar la calidad de los datos:

- **Nuevas variables:** conversión de *age* a años (*age_years*), bandas funcionales (*age_band*) y cuantiles (*age_quantile*), *BMI* y *BMI_cat* (según rango OMS).
- **Unidades validadas:**
 - *BMI*: kg/m² (rango aceptado: 10–80).
 - *ap_hi*: mmHg (presión sistólica, rango aceptado: 90–200).
 - *ap_lo*: mmHg (presión diastólica, rango aceptado: 60–120).
- **Reglas aplicadas:**
 - Eliminación de registros con *ap_lo > ap_hi* (1,77%).
 - Eliminación de valores fuera de rango clínico.
- **Flags creadas:**
 - **HTA**: *ap_lo > 80 mmHg*.
 - **Riesgo clínico**: *ap_hi ≥ 140 mmHg*, *ap_lo ≥ 90 mmHg*, *BMI ≥ 30 kg/m²*, *age_years ≥ 60 años*.
 - **Riesgo de estilo de vida**: *smoke = si*, *alco = si*, *active = no*.
- **N afectado:** Se eliminaron aproximadamente **2,72%** de los registros por winsorización.

ROC-AUC / PR-AUC (CV y test), Brier y curva de fiabilidad

Modelo	Métrica	CV	Test	Test Bootstrapping	IC 95%
Regresión Logística	ROC-AUC	$0,792 \pm 0,003$	0,787	0,787	0,779 – 0,795
	PR-AUC	$0,771 \pm 0,005$	0,766	0,766	0,755 – 0,777
	Brier score	-	0,189	-	-
	ECE	-	0,0296	-	-
Random Forest	ROC-AUC	$0,802 \pm 0,003$	0,795	0,795	0,787 – 0,802
	PR-AUC	$0,782 \pm 0,006$	0,775	0,776	0,764 – 0,787
	Brier score	-	0,184	-	-
	ECE	-	0,0127	-	-

- **Curvas ROC**: buena discriminación global.
- **Curvas PR**: precisión aceptable en contexto de clase minoritaria.
- **Test Bootstrapping**: modelos estables y robustos.
- **Calibración probabilidades**: **método isotonic** para predicciones con precisión.
- **Brier score**: aceptable, pero no óptimo.
- **ECE**: probabilidades predichas reflejan muy bien el riesgo. La calibración isotónica ha sido efectiva.



- **Curva de fiabilidad:** modelo RF no solo tiene un buen Brier, sino que no sobreestima ni subestima el riesgo en ningún bin. Cuidado en los extremos.
- **Modelo escogido:** Random Forest

Umbral elegido y justificación (Youden / coste 5:1 / top-k)

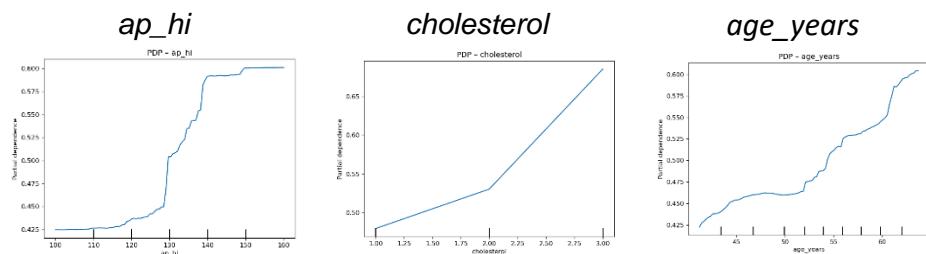
Modelo	Criterio	Umbral	Sens	Espec	PPV	NPV	F1
Regresión Logística	Youden J	0,477	0,693	0,752	0,732	0,714	0,712
	Coste 5:1	0,150	0,986	0,114	0,521	0,893	0,682
	Top-k	0,751	0,341	0,938	0,843	0,592	0,485
Random Forest	Youden J	0,466	0,707	0,748	0,733	0,723	0,720
	Coste 5:1	0,170	0,989	0,110	0,521	0,912	0,683
	Top-k	0,806	0,342	0,940	0,847	0,593	0,488

RF + Youden J → mejores métricas, más beneficio en la mayoría de los puntos en el DCA (0,2 – 0,6) y mejor umbral operativo ya que muestra mejor equilibrio.

PI + PDP/ALE (2–4 gráficos) con interpretación clínica.

PI: variables más relevantes: *ap_hi* (0, 1155), *cholesterol* (0, 0405), *age_years* (0, 01239).

PDP: cambios de tendencia según clínica en todas las variables.



ALE: *BMI* riesgo en bajo peso y sobrepeso. En *ap_hi* mayor riesgo clínico cuando es alta. En *ap_lo* hay menor riesgo clínico cuando es alta sin umbral clínico claro.

Las interacciones indican que:

- BMI alto ($\geq 35 \text{ Kg/m}^2$) y ap_hi elevada ($\geq 140 \text{ mmHg}$) → mayor riesgo.
- BMI bajo ($< 20 \text{ Kg/ m}^2$) y ap_hi normal ($< 120 \text{ mmHg}$) → riesgo elevado.
- Age avanzada (> 60 años) + ap_hi elevada ($> 140 \text{ mmHg}$) → efecto sinérgico → riesgo amplificado.
- Age avanzada (> 60 años) + ap_lo baja ($< 70 \text{ mmHg}$) → efecto sinérgico → riesgo elevado.

Subgrupos (sexo/edad) y comentarios de equidad

- Sexo: equidad por sexo, con métricas casi idénticas entre hombres y mujeres.
- Edad: variabilidad por edad → mayor sensibilidad en adultos mayores pero menor especificidad, y lo contrario en jóvenes.
- Sesgo por edad y riesgo clínico donde hay más positivos en mayores y en pacientes con HTA o riesgo alto.
- Alta sensibilidad en grupos clínicamente relevantes, pero con pérdida de especificidad.

Límites y plan de siguientes pasos

- Límites:
 - No hay seguimiento longitudinal ni validación temporal o geográfica.
 - Variables importantes están discretizadas (*colesterol, gluc*).
 - Calibración subóptima en los extremos (DCA válido entre 0,2- 0,7).
 - Revisar distribución por edad, BMI y presión arterial ya que las métricas muestran disparidades.
- Siguientes pasos generales:
 - Validación externa en otras cohortes clínicas.
 - Simulación de impacto clínico con costes reales.
 - Monitorizar periódicamente las variables clave y la tasa de positivos para detectar cambios poblacionales.
- Siguientes pasos del modelo:
 - Respecto a los extremos hacer una calibración local y visualizar mejor la incertidumbre.
 - Recalibración por subgrupo (age/BMI/ap): curvas de fiabilidad, cambios umbral, validad sens y PPV por grupo.
 - SHAP por paciente, interacciones clínicas más complejas y visualización de zonas de incertidumbre.
 - Integrar el modelo en un sistema de priorización, no como clasificador binario.