



Bank Account Fraud: Solución

EQUIPO 5

Carlos Alberto Mentado Reyes A01276065

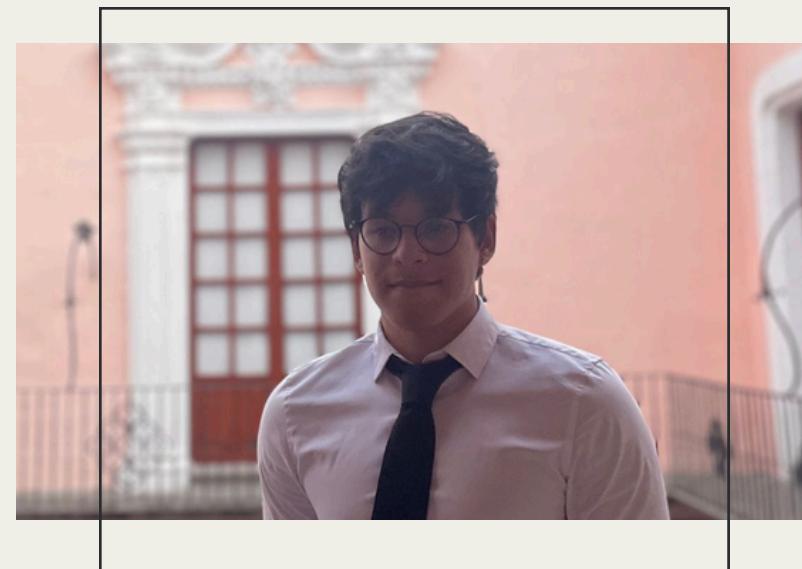
Raymundo Ivan Diaz Alejandre A01735644

José Eduardo Puentes Martínez A01733177

Fernanda Díaz Gutiérrez A01639572



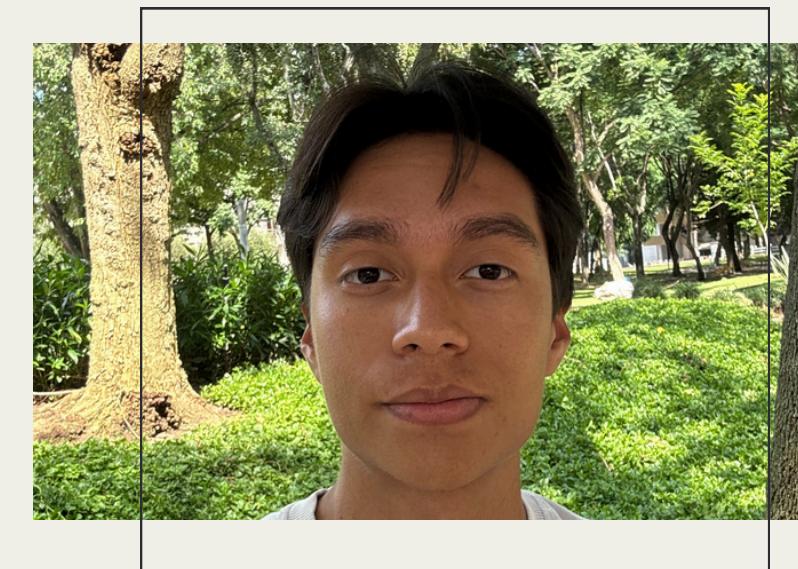
EL EQUIPO



**Raymundo Ivan Diaz
Alejandre**



**Carlos Alberto
Mentado Reyes**



**José Eduardo
Puentes Martínez**



**Fernanda Diaz
Gutiérrez**

Introducción al problema



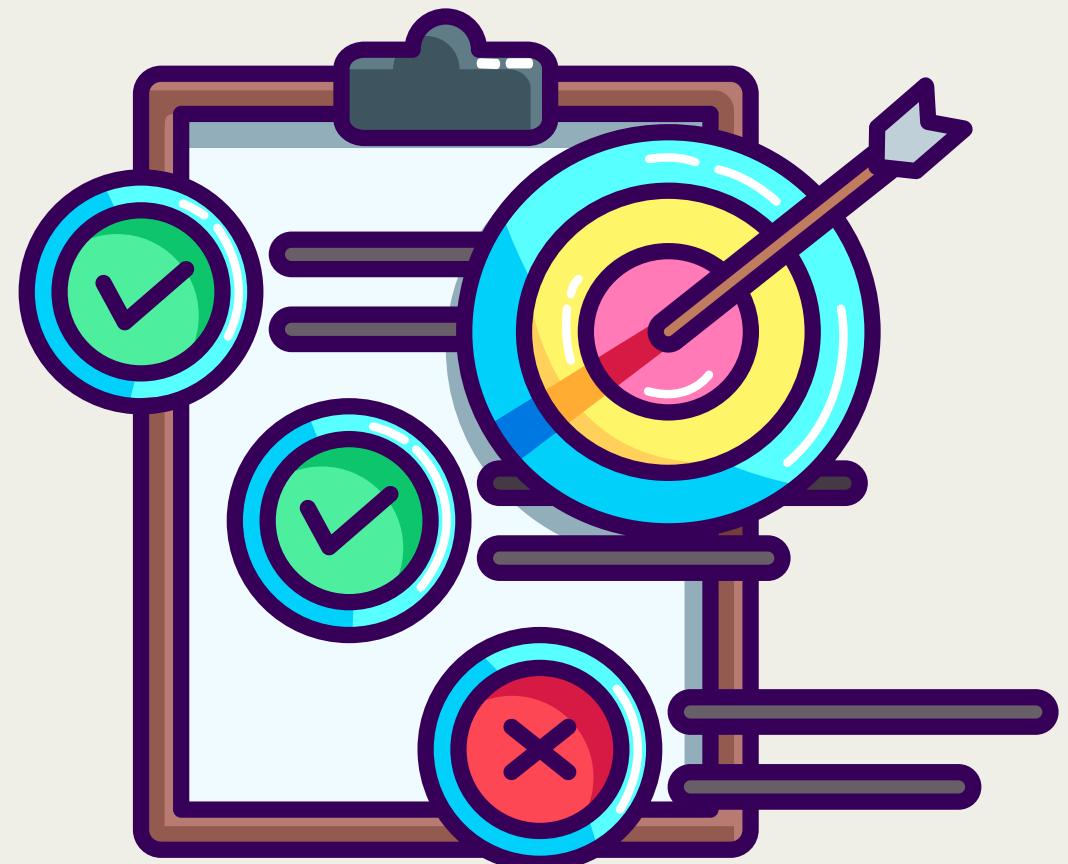
EXPLICANDO EL PROBLEMA

BAF consta en una serie de 6 datasets sintéticos que representan un reto realista basado en fraudes de cuentas de banco para el entrenamiento y validación de modelos de clasificación



OBJETIVO DEL RETO

- Obtener el mayor recall posible mientras FPR se mantiene en 5%
- Basar los resultados en una curva ROC
- Mantener imparcialidad



RETOS PRESENTADOS

- Desbalanceo de clase
- Variables categóricas
- Atributos protegidos
- Falta de valores representado como -1
- Sesgos presentes:
 - Disparidad de grupos
 - Disparidad de prevalencia
 - Disparidad en separabilidad



DATASETS

Cada dataset cuenta con 1 millón de entradas y 31 (o 33) columnas de características y una variable objetivo

Dataset	Description
Base	Sampled to best represent original dataset.
Variant I	Has <u>higher group size disparity</u> than base.
Variant II	Has <u>higher prevalence disparity</u> than base.
Variant III	Has <u>better separability</u> for one of the groups.
Variant IV	Has <u>higher prevalence disparity in train</u> .
Variant V	Has <u>better separability in train</u> for one of the groups.

Contenido Técnico

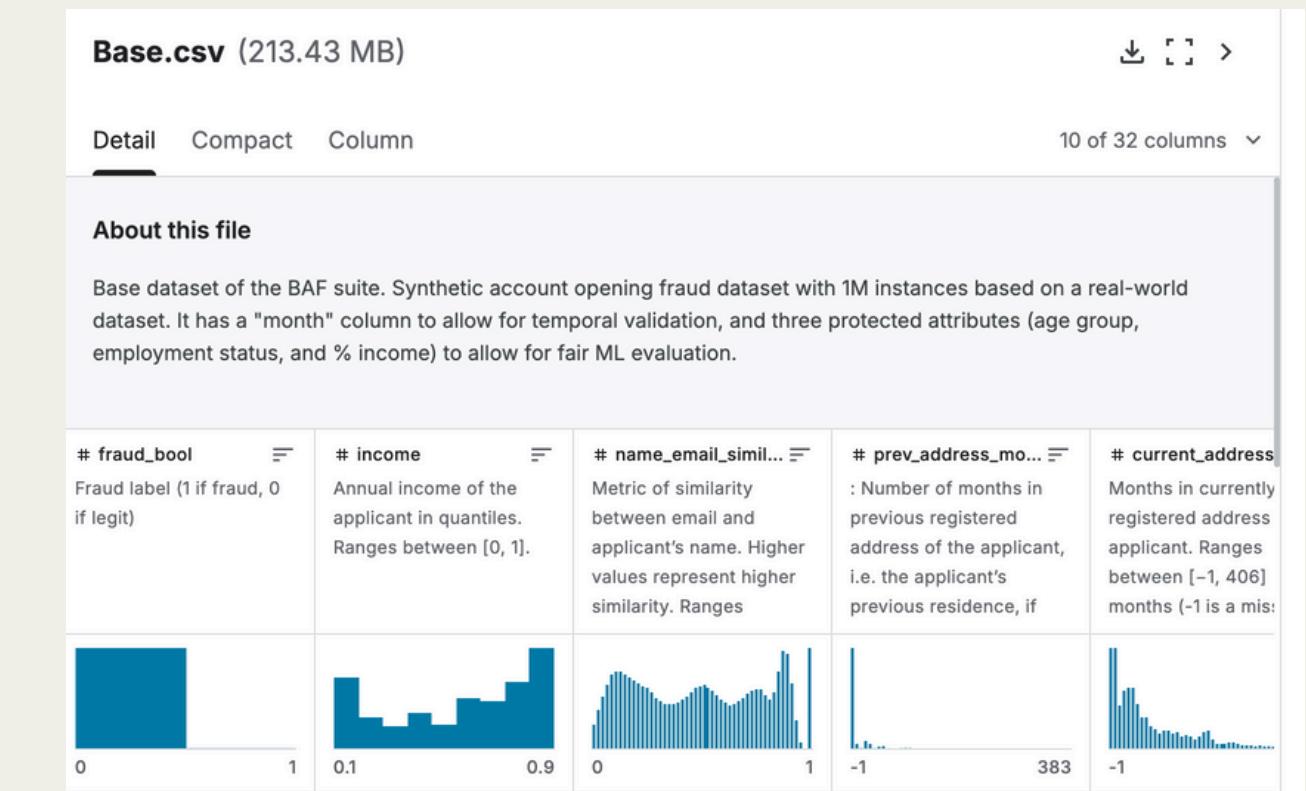


DATOS: BASE

- Base.csv: 1,000,000 registros
- 31 características (numéricas y categóricas) + 1 variable objetivo
- Variable target: **fraud_bool** (1 = fraude, 0 = no fraude)
- Desbalance fuerte: ≈ 0.8–1% fraudes

RETOS

- Desbalance severo (=99:1)
- Missing codificados como -1 en varias columnas (bank_months_count, prev_address_months_count, session_length_in_minutes, device_distinct_emails_8w)
- Mixto: numéricas + categóricas
- Posibles sesgos (edad, ingreso, empleo)



DATOS: LIMPIEZA Y TRANSFORMACIÓN

DE -1 A NAN SIN FUGAS

- Reglas: si valor < 0 → NaN
- Se aplica en train/val/test (mismo criterio)
- Para que los imputers trabajen con información correcta y no con un -1.

TRANSFORMACIÓN

- **Numéricas:** Imputación mediana → StandardScaler
- **Categóricas:** Imputación moda → One Hot Encoding
- Fit SOLO en train → aplicar en val/test



DATOS: LIMPIEZA Y TRANSFORMACIÓN

SPLIT TEMPORAL

- Train: months 0–5 | Validation: 6 | Test: 7
- Respeta la direccionalidad temporal
- Early stopping en validation

0 - 5 Train

6 validation

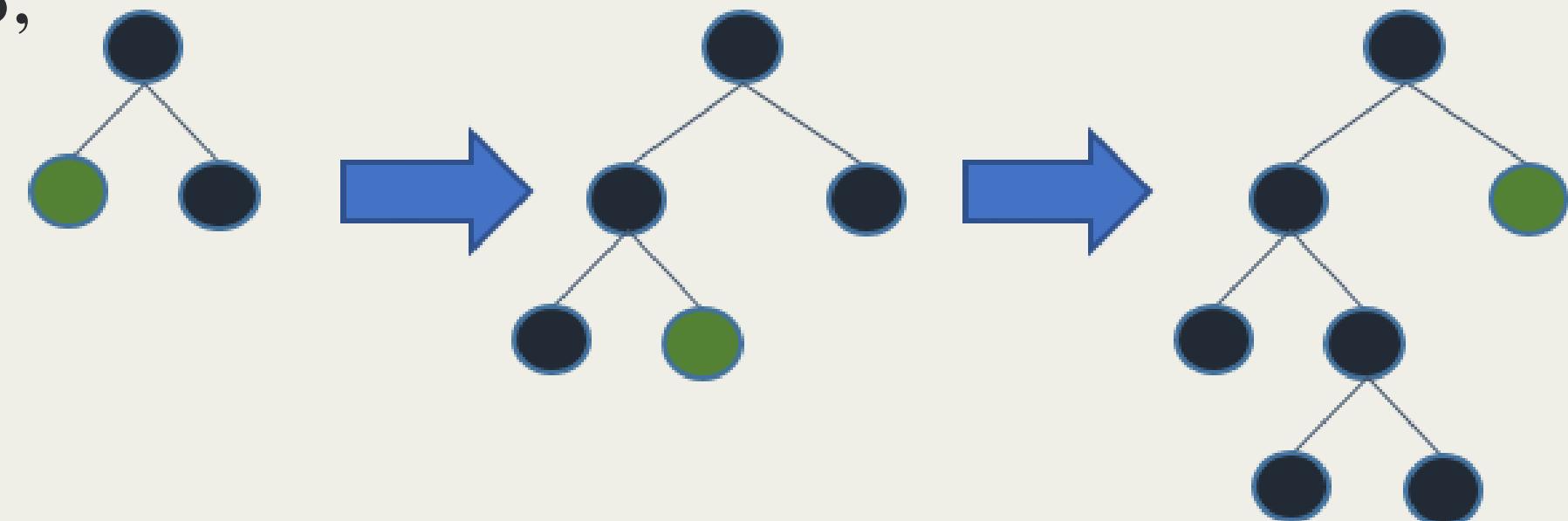
7 test

MODELO (LIGHTGBM)

- LightGBM (GBDT) boosting basado en árboles de decisiones para tabulares: rápido, robusto, maneja no linealidades/altas cardinalidades

HIPERPARÁMETROS:

- n_estimators=1800, learning_rate=0.03
- num_leaves=128, min_child_samples=80
- subsample=0.8, colsample_bytree=0.8, reg_lambda=1.0
- **class_weight**: pos = neg/pos
- EarlyStopping(200) con AUC en val



MODELO (LIGHTGBM)

MÉTRICAS Y POR QUÉ

- **Primarias:**
- Recall en $FPR=5\%$ (métrica oficial, maximizar fraudes detectados sin pasar 5% de falsos positivos)
- AUC, AP (PR-AUC, Avg Precision) para ranking/desbalance
- **Fairness (método):** Predictive Equality = comparar FPR entre grupos

FIJANDO EL UMBRAL

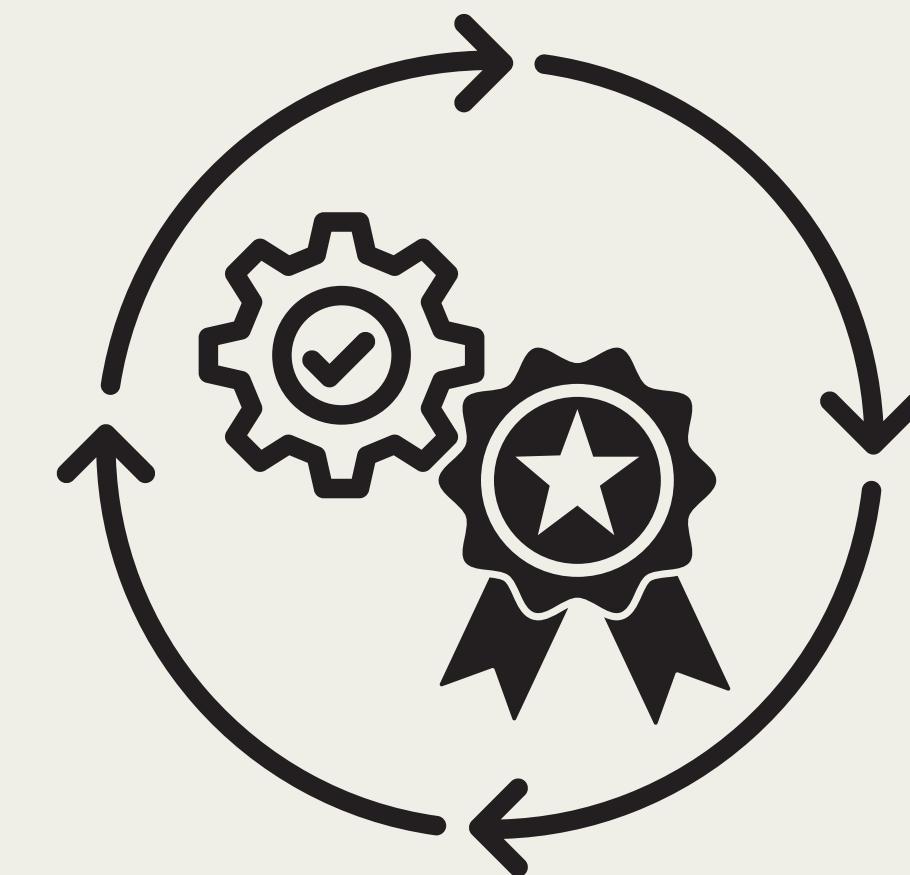
1. Obtener scores p_val en validación
2. Tomar solo negativos ($y=0$)
3. Calcular el percentil 95 de esos scores
4. Ese valor = umbral operativo (deja 5% de negativos arriba)
5. Congelar umbral y evaluar en test



MODELO Y ESTADÍSTICA

ROBUSTEZ Y FAIRNESS

- **Sanity checks:** distribución de scores, estabilidad AUC (val vs test), análisis de outliers
- AUC estable entre validación y prueba, lo que muestra que no hubo sobreajuste.
- Distribución de probabilidades para confirmar que el modelo realmente separa pos/neg.
- **Fairness (Predictive Equality):**
- Definición de grupos: $\text{age} < 50$ vs ≥ 50
- FPR por grupo y ratio min/max
- ¿Qué tan equitivamente trata el modelo a distintos grupos de edad?



Resultados



Validación

AUC: **0.890**

AP (PR-AUC): **0.176**

Umbral operativo (FPR 5%): **0.768**

FPR: **0.050**

Recall: **0.524**

Test (usando umbral de validación)

AUC: **0.890**

AP (PR-AUC): **0.211**

FPR (umbral val): **0.042**

Recall (umbral val): **0.535**

Recall (curve, FPR=5%): **0.562**

Fairness: Predictive Equality (edad)

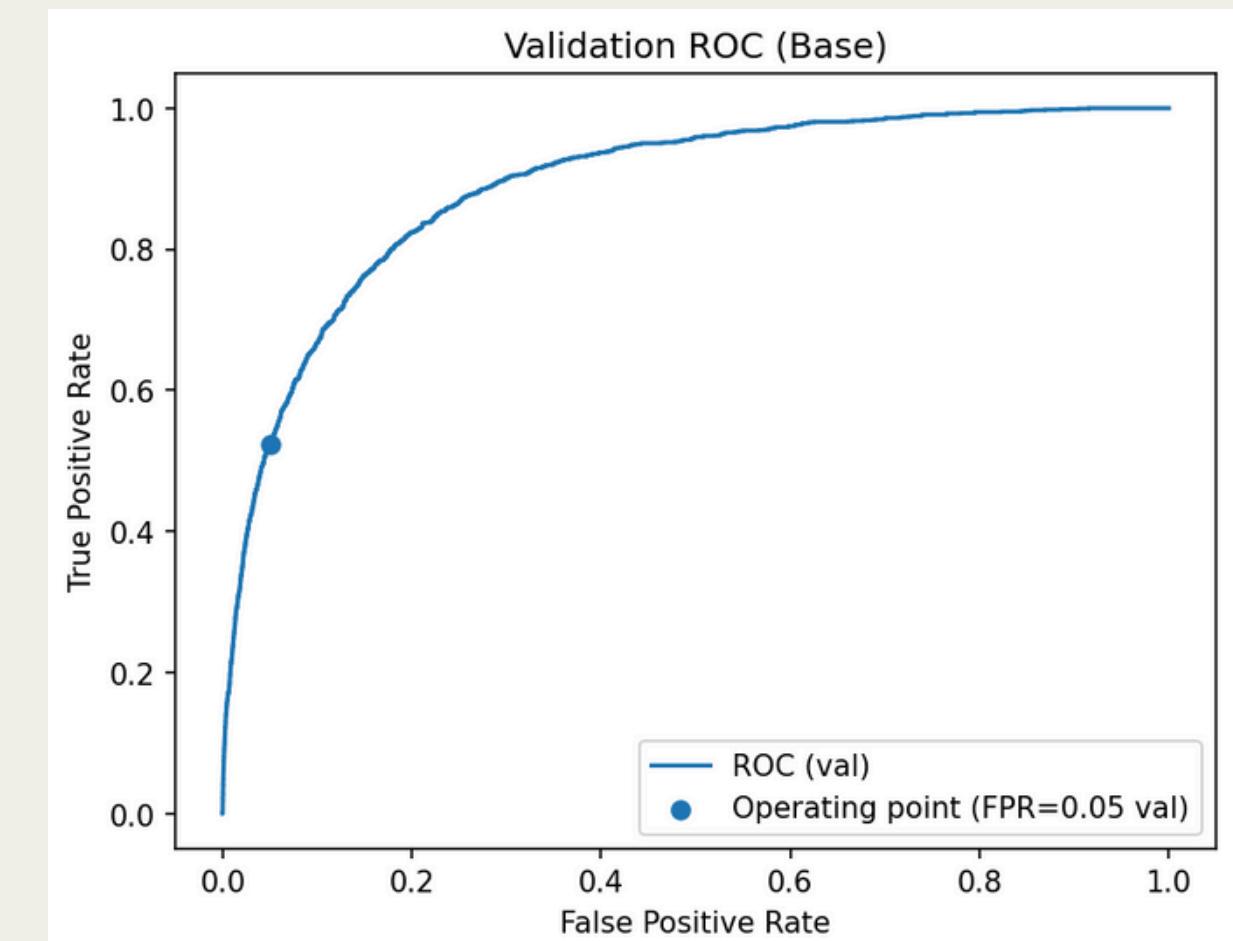
FPR <50 años: **0.033**

FPR ≥50 años: **0.101**

Ratio FPR (min/max): **0.329**

ROC VALIDACIÓN

- AUC = 0.89, buena capacidad de discriminación.
- Umbral operativo de 0.768, lo cual asegura un 5% de FPR
- El modela recupera un 52% de los fraudes.



RESULTADOS EN TEST

- Aplicando el umbral al test
 - AUC se mantiene
 - AP mejora
 - NO hay sobreajuste
 - FPR baja 4.2%
 - Recall sube 53%
- Si ajustamos umbral con
 - FPR de 5%
 - Recall subiría a 56%

IMPARCIALIDAD POR EDAD

- Predictive Equality
- Edad
 - > 50 , FPR 10%
 - < 50 , FPR 3%
- Ratio 0.329
- Adultos mayores reciben más falsos positivos

Detección de fraudes

AUC (test)

0.890

LightGBM (operating @ 5% FPR val)

Recall@test-thr (test)

53.5%

FPR@test-thr (test)

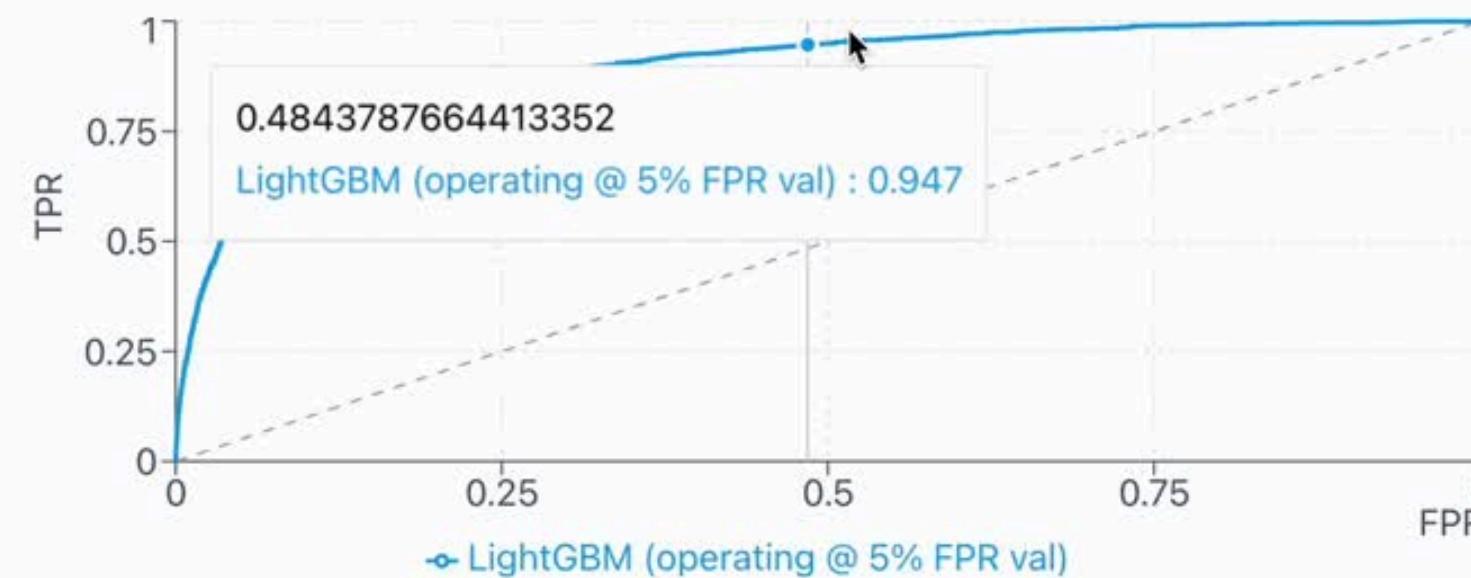
4.2%

AP (test)

0.211

ROC Curve

AUC = 0.890



Punto operativo (validación) & métricas

Fairness · Predictive Equality (FPR por edad)

Base

Detección de fraudes

AUC (test)

0.890

LightGBM (operating @ 5% FPR val)

Recall@test-thr (test)

50.0%

FPR@test-thr (test)

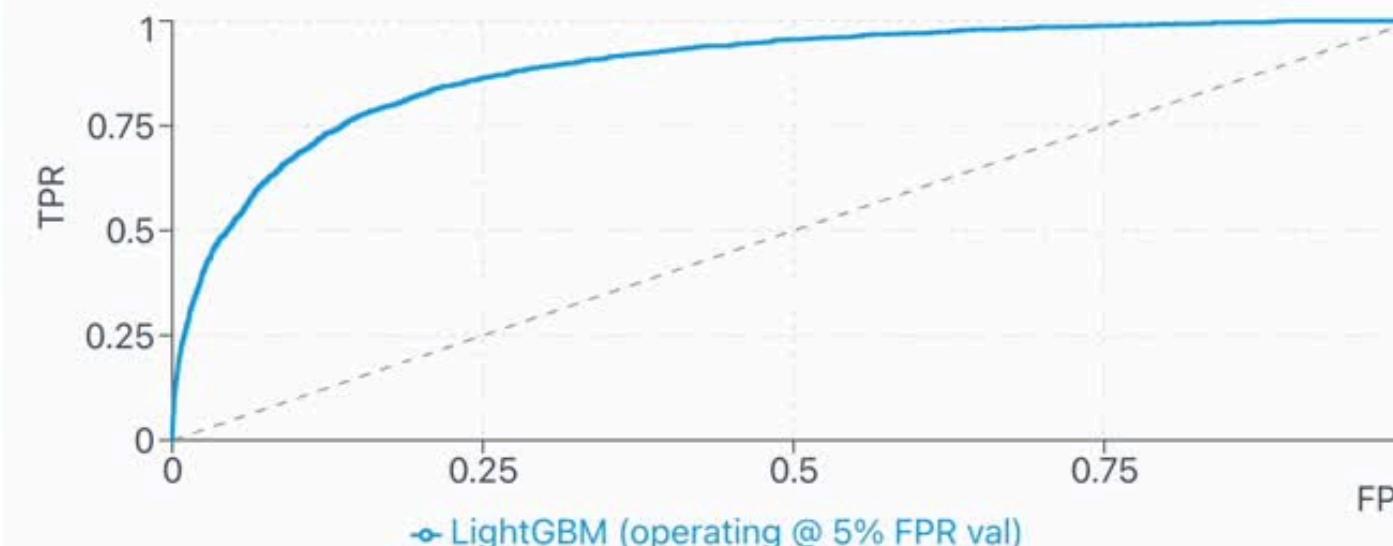
4.5%

AP (test)

0.215

ROC Curve

AUC = 0.890



Punto operativo (validación) & métricas

Umbral @ FPR 5% (val)

FPR (val)

Fairness · Predictive Equality (FPR por edad)



Variante 1

Detección de fraudes

AUC (test)

0.895

LightGBM (operating @ 5% FPR val)

Recall@test-thr (test)

53.2%

FPR@test-thr (test)

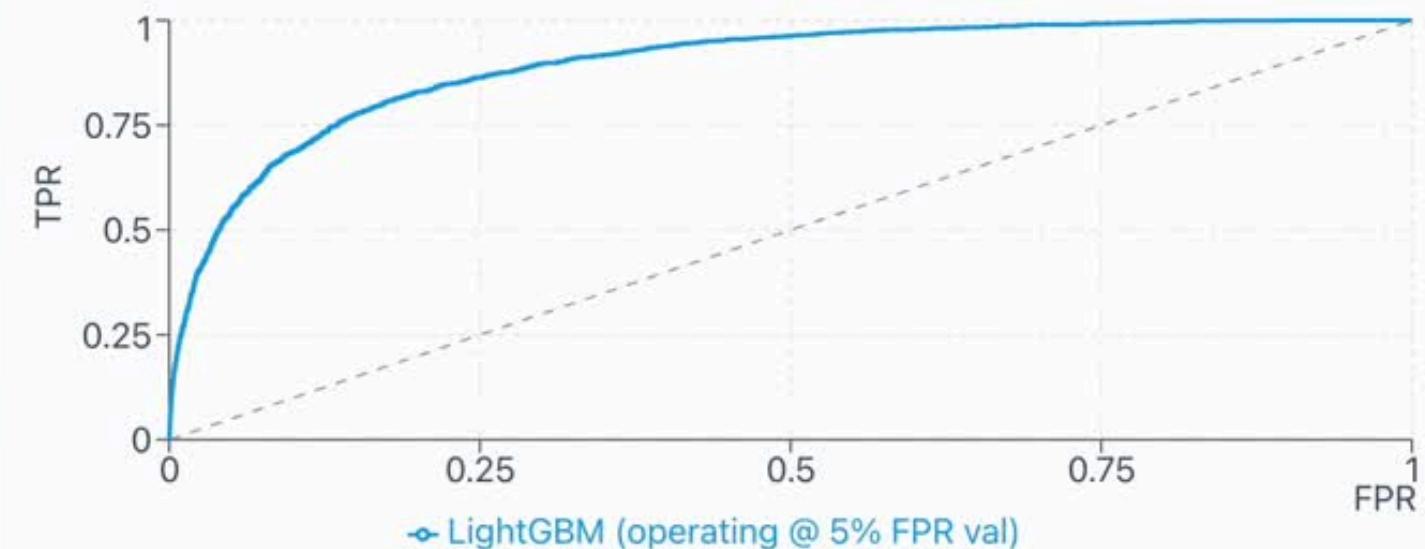
4.7%

AP (test)

0.208

ROC Curve

AUC = 0.895



Punto operativo (validación) & métricas

Umbral @ FPR 5% (val)

0.777

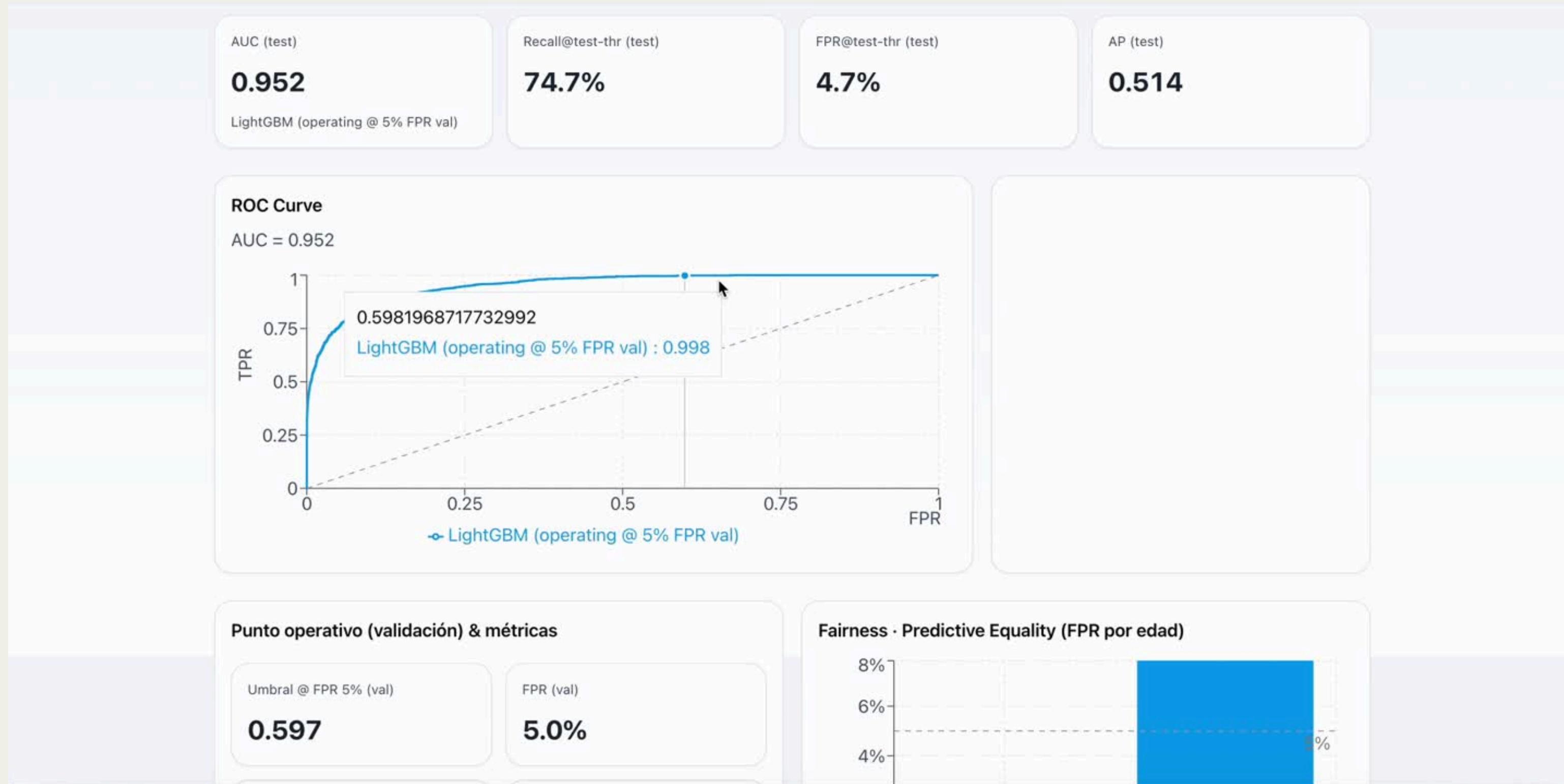
FPR (val)

5.0%

Fairness · Predictive Equality (FPR por edad)

8%

Variante 2



Variante 3

Detección de fraudes

AUC (test)

↑0.855

LightGBM (operating @ 5% FPR val)

Recall@test-thr (test)

42.2%

FPR@test-thr (test)

4.5%

AP (test)

0.095

ROC Curve

AUC = 0.855



Punto operativo (validación) & métricas

Umbral @ FPR 5% (val)

0.691

FPR (val)

5.0%

Fairness - Predictive Equality (FPR por edad)

8%
6%

Variante 4

Detección de fraudes

AUC (test)

0.768

LightGBM (operating @ 5% FPR val)

Recall@test-thr (test)

27.0%

FPR@test-thr (test)

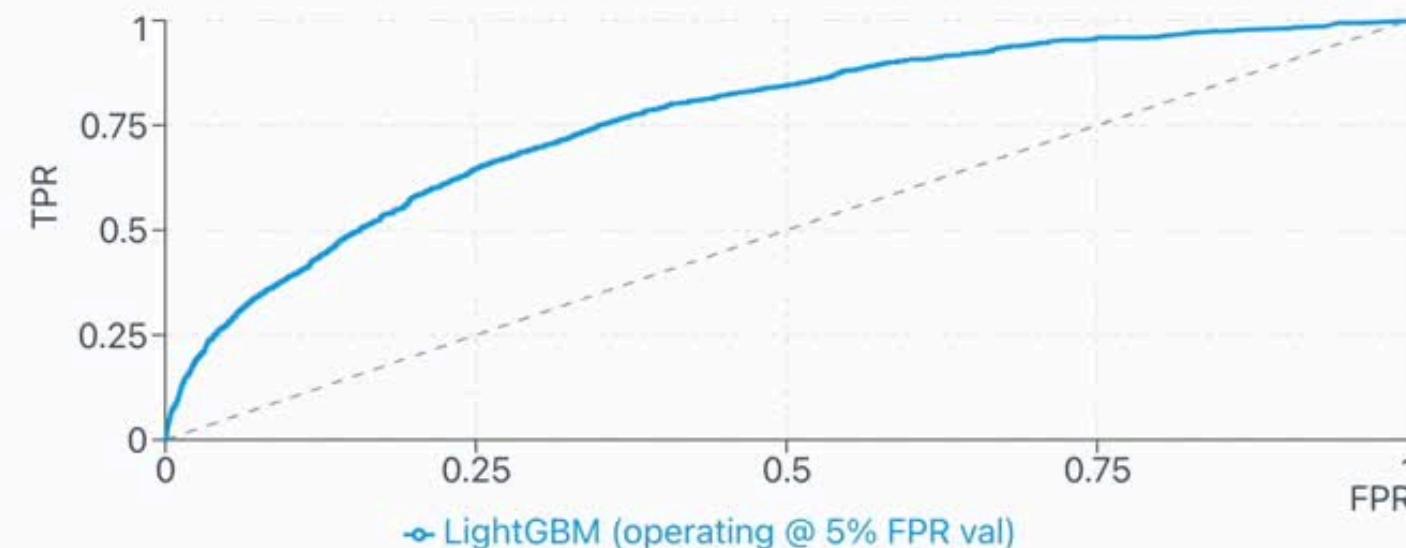
4.7%

AP (test)

0.041

ROC Curve

AUC = 0.768

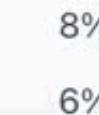


Punto operativo (validación) & métricas

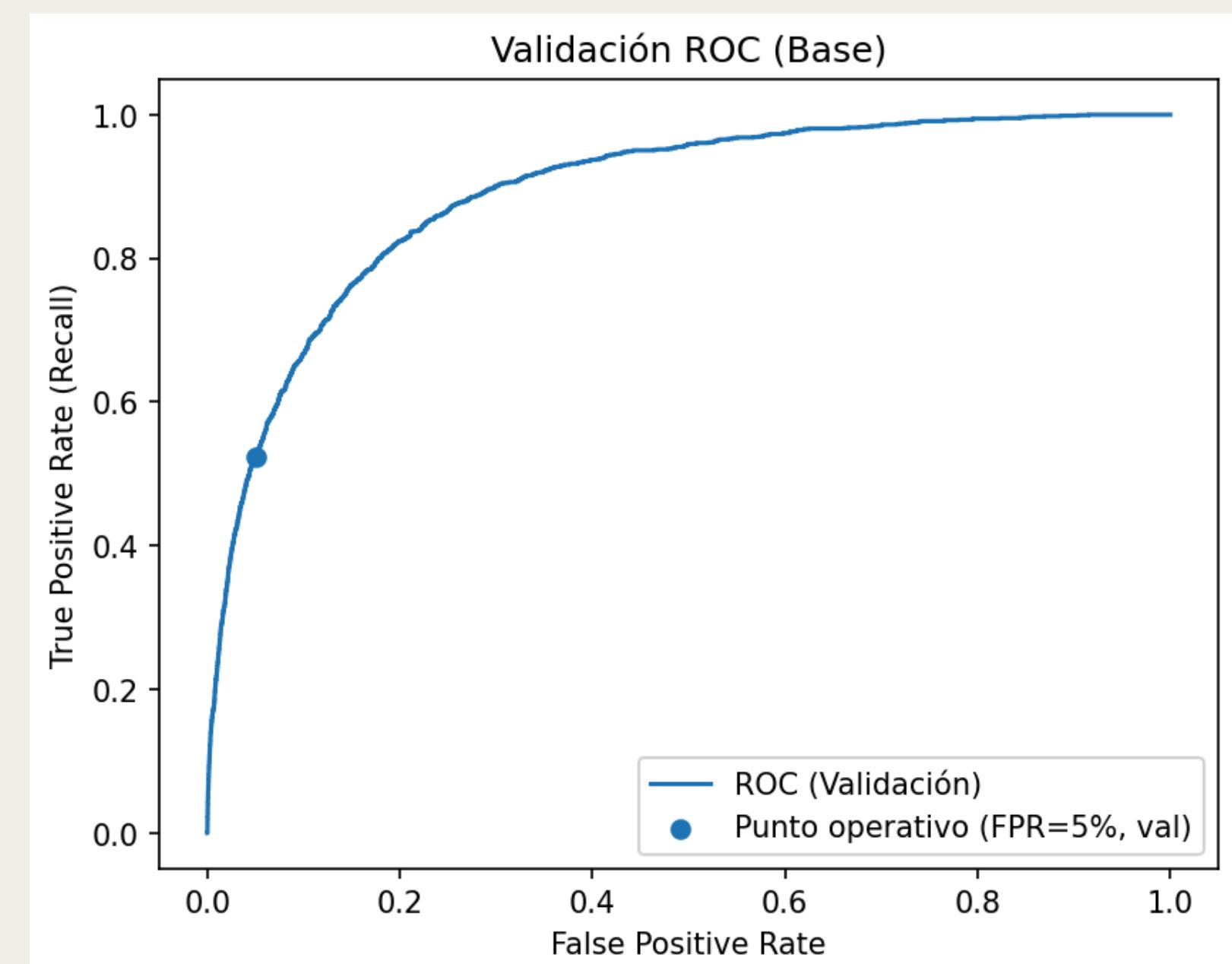
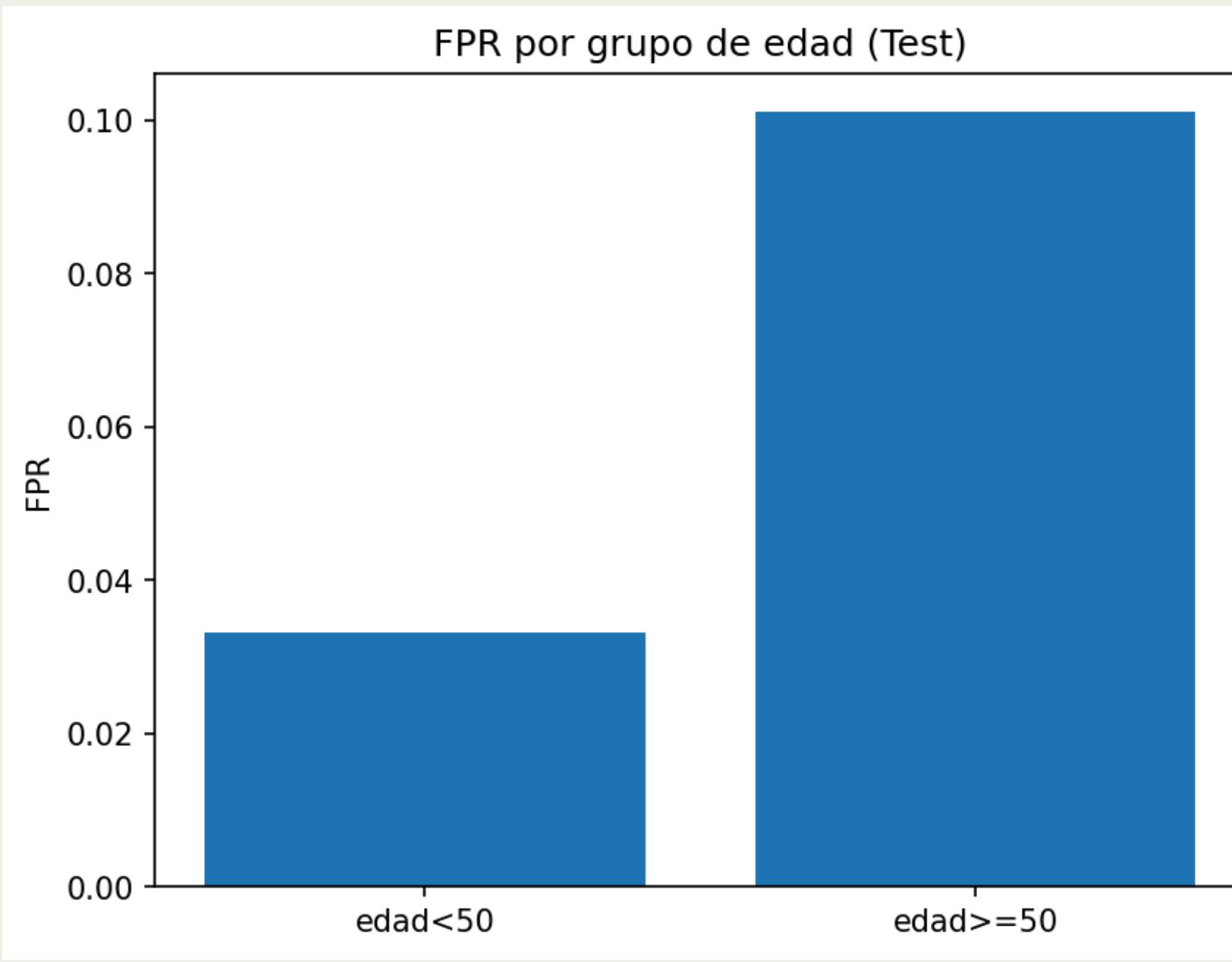
Umbral @ FPR 5% (val)

FPR (val)

Fairness · Predictive Equality (FPR por edad)



Variante 5



CONCLUSIONES



Detectamos

56% de fraudes al
5% de FPR



AUC y AP altos

El modelo analiza
muy bien, incluso
teniendo un alto
desbalance



No hay overfitting

AUC y AP son
consistentes entre
val y test



Fairness

Se detecto un sesgo
esperado, en este
caso siendo mas
falsos positivos en
mayores de 50 años

¡Gracias!

[HTTPS://GITHUB.COM/CYBERPEONY/BAF_IA_AVANZADA_E5](https://github.com/cyberpeony/BAF_IA_AVANZADA_E5)

