

Análisis de Finanzas y Riesgo Crediticio

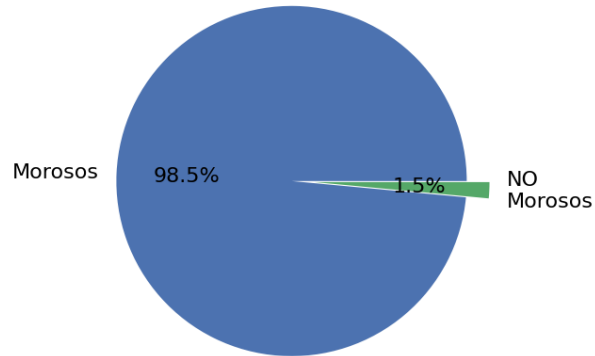
- ¿Qué umbrales de saldo podrían indicar mayor riesgo de morosidad?



ANÁLISIS EXPLORATORIO

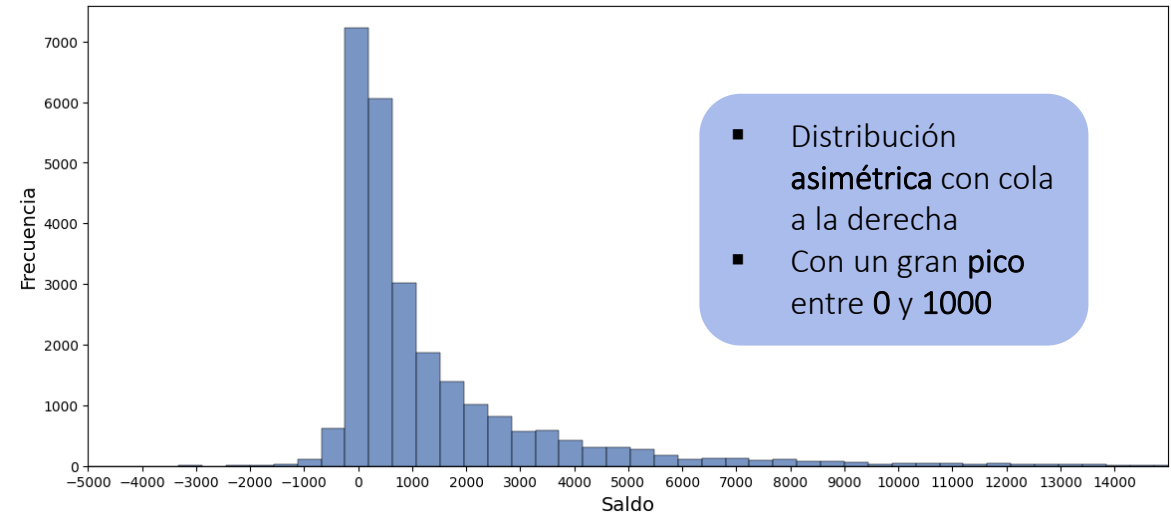
Examinamos las variables “default” y “balance”

Distribución de Morosidad



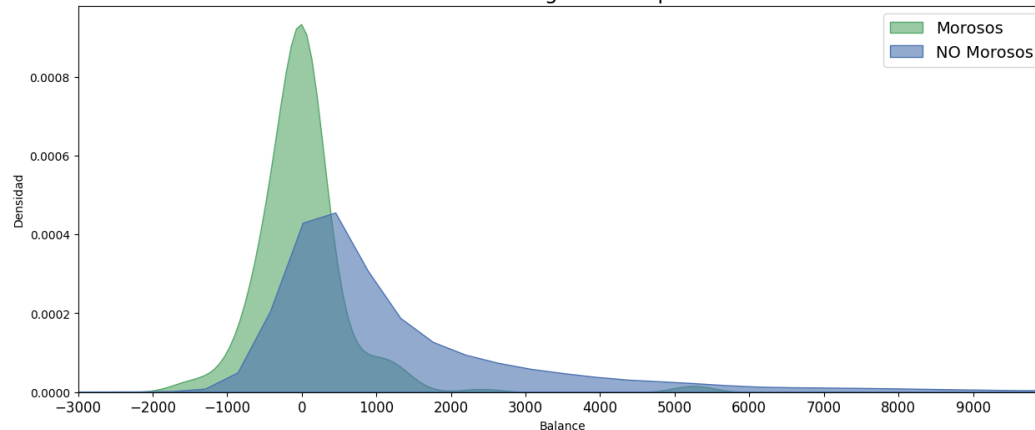
- Distribución **muy asimétrica**
- El **1,5%** de los clientes en estado de morosidad

Distribución del Saldo de los Clientes



- Distribución **asimétrica** con cola a la derecha
- Con un gran **pico** entre 0 y 1000

Densidad de Saldo segmentado por Morosidad



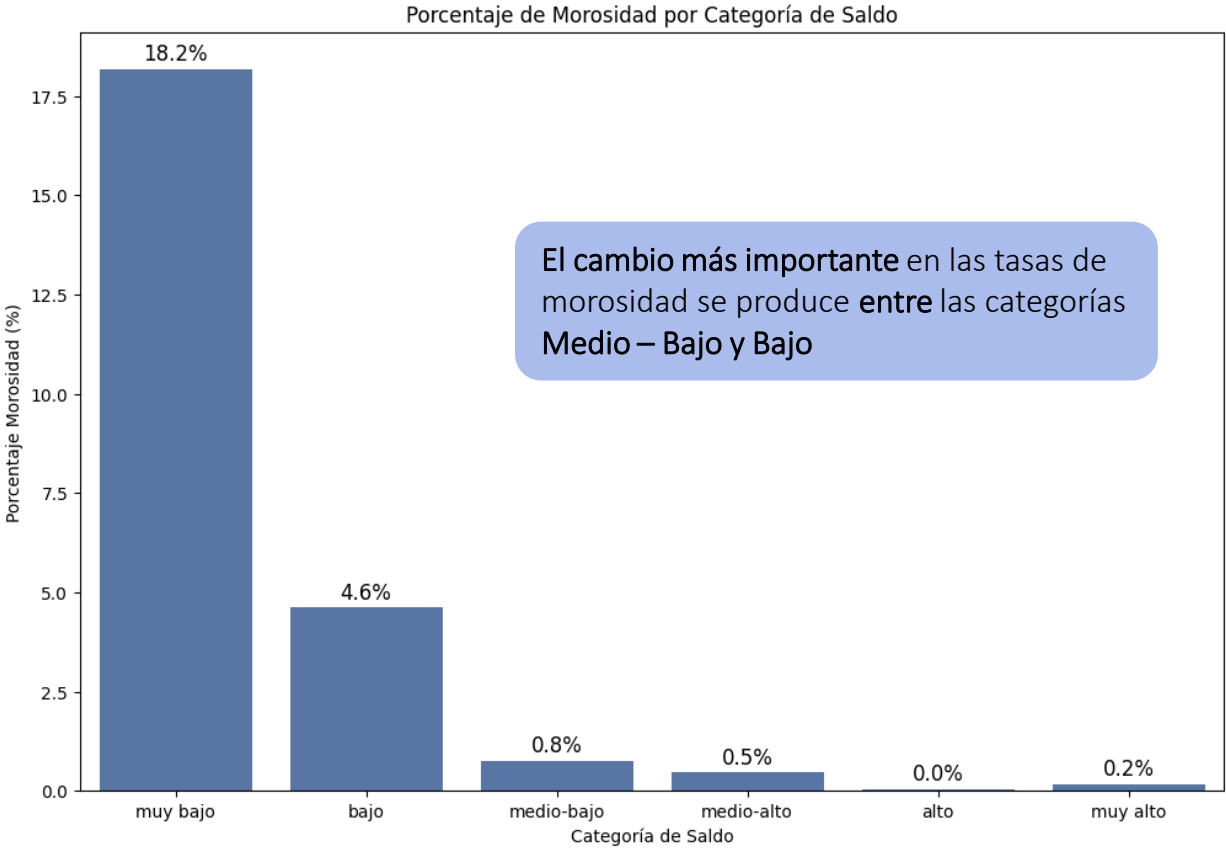
- Ambas presentan **distribuciones normales**.
- La distribución de **morosos** tiene un mayor **pico** cerca del 0.
- **No morosos** tiene una **cola** más larga.

CATEGORIZACIÓN DE LAS VARIABLES

Categorizamos por Saldo para ver como varia la tasa de morosidad en cada una.

Categoría	Saldo mínimo	Cantidad de Clientes
Muy Bajo	-6847	11
Bajo	-2049	6506
Medio Bajo	123	6564
Medio Alto	543	6535
Alto	1704	4081
Muy Alto	4079	2466

La gran mayoría de los clientes están en las 4 categorías centrales

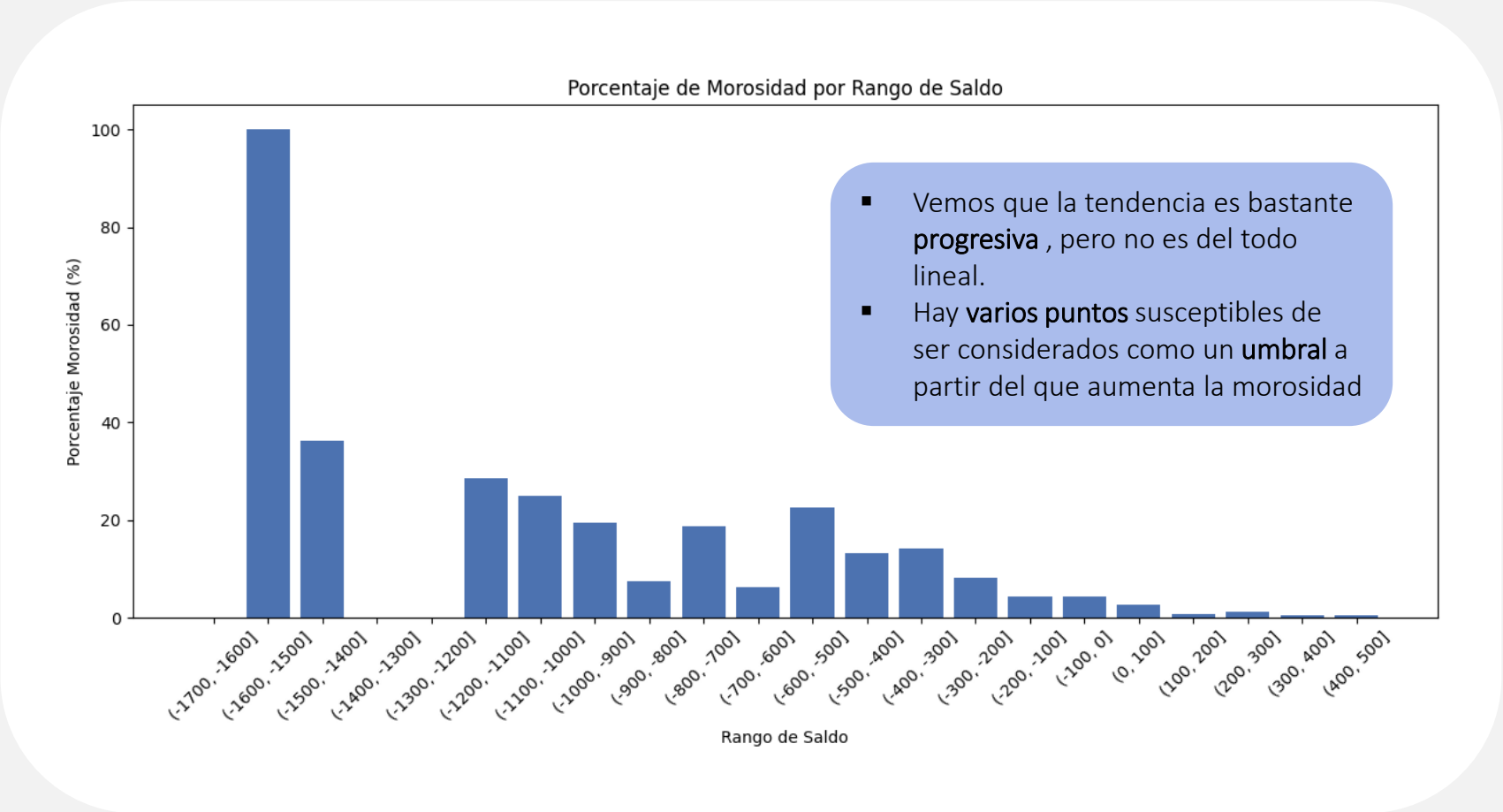


También observamos un pequeño número de valores muy extremos que dificultan el análisis por lo que decidimos continuar sin estos **Outliers** aplicando el método $IQR \cdot 1.5$. Así podremos **centrarnos en la parte** de la muestra **donde se producen los cambios** de tendencia para buscar el **umbral**.

CATEGORIZACIÓN DE LAS VARIABLES

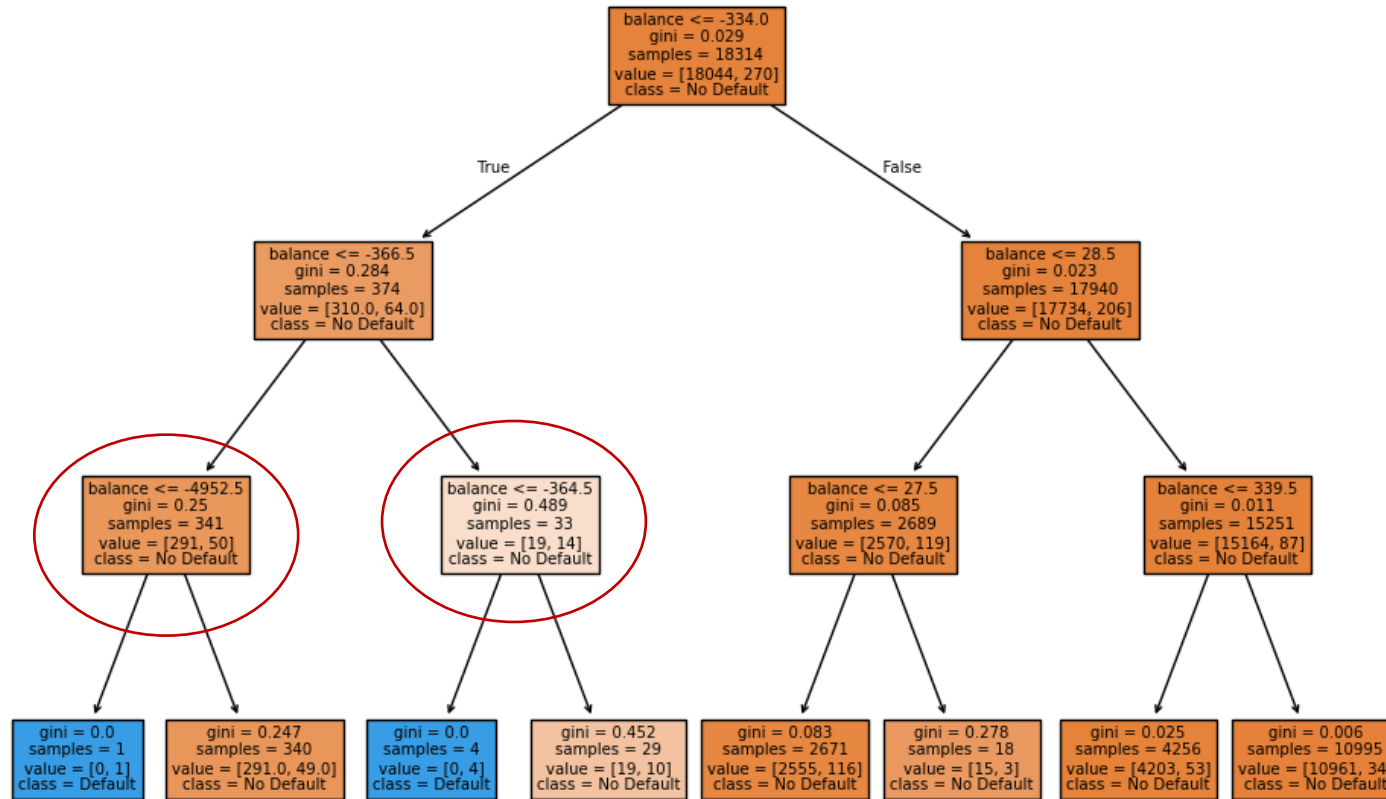
Hacemos categorías más pequeñas y nos centramos en el rango de Saldo donde ocurren los cambios más significativos

	rango	porcentaje morosidad (%)	diferencia (%)
21	(400, 500]	0.580431	0.000000
20	(300, 400]	0.639795	10.227584
19	(200, 300]	1.251422	95.597270
18	(100, 200]	0.784698	-37.295466
17	(0, 100]	2.696544	243.640809
16	(-100, 0]	4.276169	58.579629
15	(-200, -100]	4.330709	1.275427
14	(-300, -200]	8.189655	89.106583
13	(-400, -300]	14.285714	74.436090
12	(-500, -400]	13.235294	-7.352941
11	(-600, -500]	22.666667	71.259259
10	(-700, -600]	6.250000	-72.426471
9	(-800, -700]	18.867925	201.886792
8	(-900, -800]	7.407407	-60.740741
7	(-1000, -900]	19.354839	161.290323
6	(-1100, -1000]	25.000000	29.166667
5	(-1200, -1100]	28.571429	14.285714
4	(-1300, -1200]	0.000000	-100.000000
3	(-1400, -1300]	0.000000	0.000000
2	(-1500, -1400]	36.363636	inf
1	(-1600, -1500]	100.000000	175.000000
0	(-1700, -1600]	0.000000	-100.000000



RELACIÓN ENTRE VARIABLES: MODELOS PREDICTIVOS

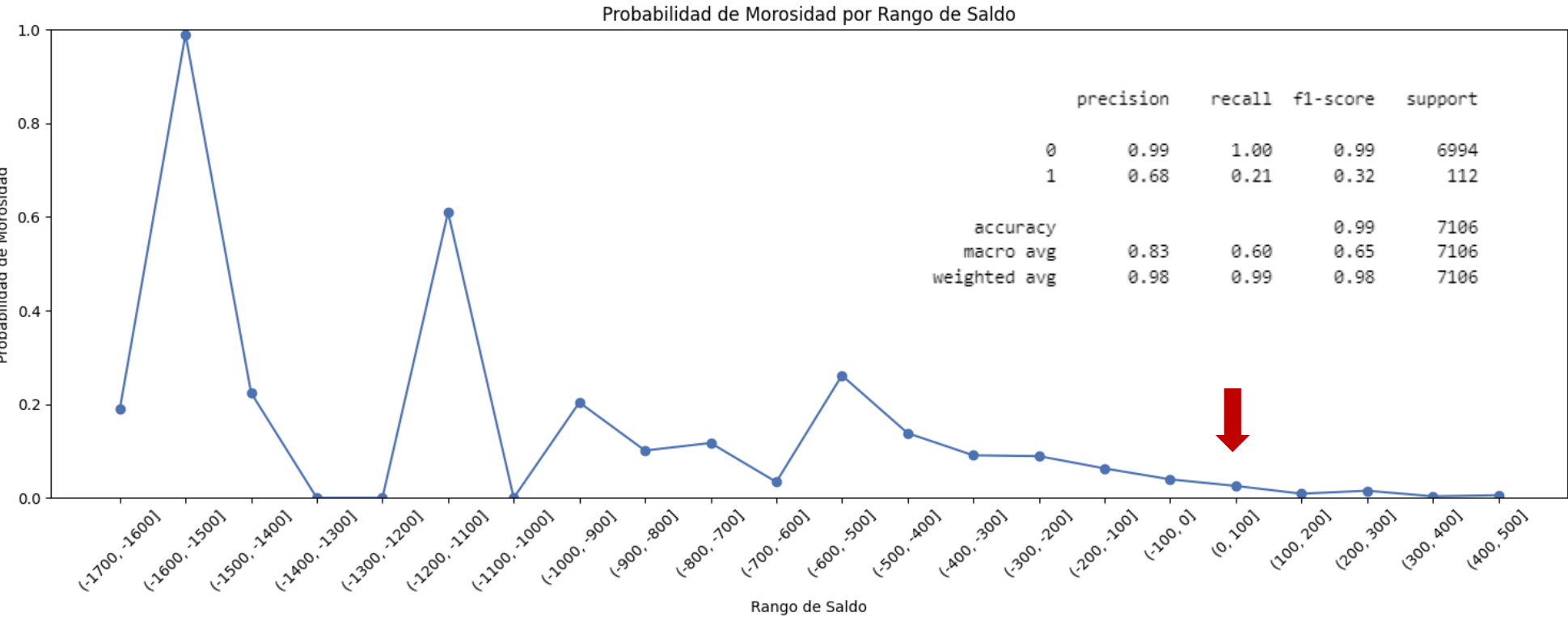
ÁRBOL DE DECISIONES



Según este modelo de clasificación, el punto donde es más fácil separar los clientes morosos de los que no lo son, es en los que tienen un saldo de $\geq -364,5$

RELACIÓN ENTRE VARIABLES: MODELOS PREDICTIVOS

RANDOM FOREST

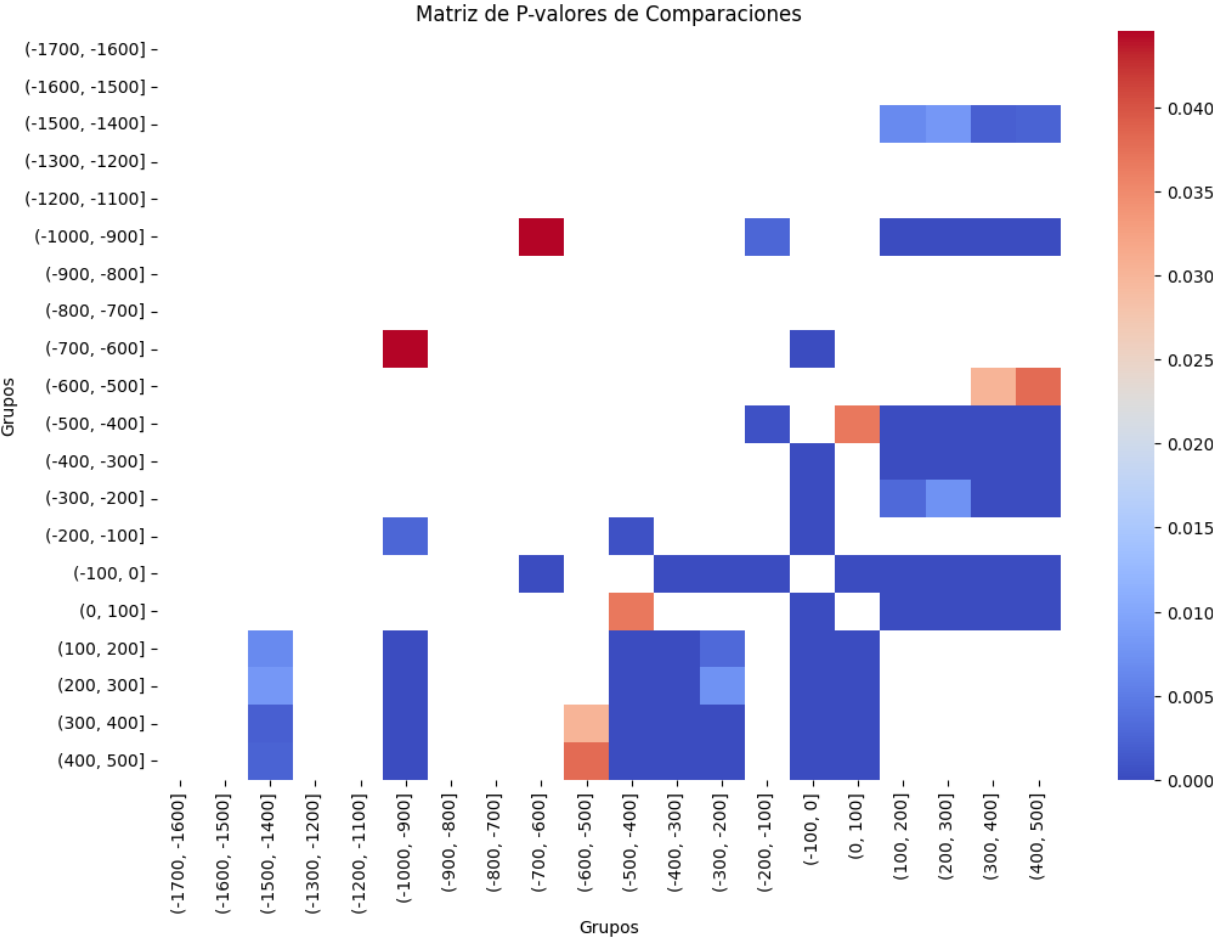


Este modelo muestra que **el punto de inflexión** en el que la **probabilidad** empieza a **crecer**, y ya casi no dejará de hacerlo, es la categoría **entre 0 y 100**



RELACIÓN ENTRE VARIABLES: MODELOS PREDICTIVOS

RANDOM FOREST: PRUEBA DE PROPORCIONES Z-TEST



Estadístico H: 1487.303456346061

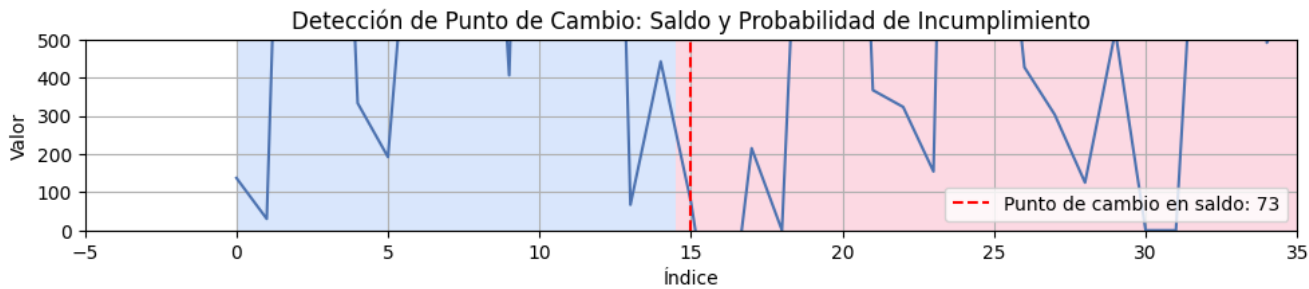
P-valor: 2.350565297963844e-304

El p-valor es menor que 0.05 podemos rechazar la hipótesis nula, lo que sugiere que los rangos de saldo influyen significativamente en la probabilidad de incumplimiento.

Vemos que los puntos donde se concentran P-Valores más bajos son de entre -100 y 100 y entre -500 y -200. Por tanto, en estos rangos es donde la prueba indica que se producen mayores cambios de probabilidad.

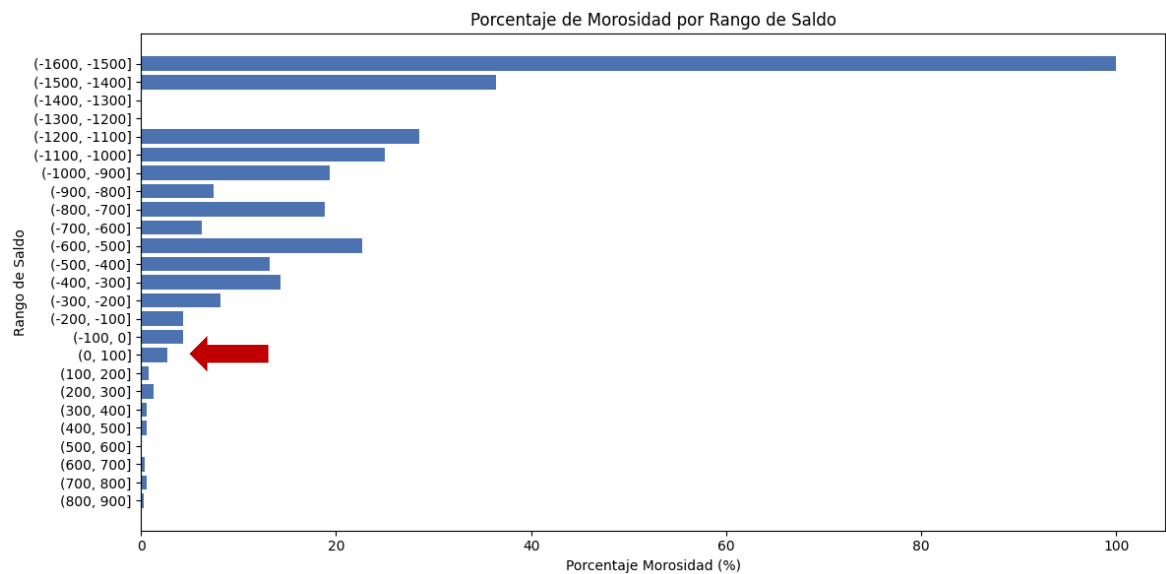
CONCLUSIÓN: ANALISIS VISUAL + MODELO PREDICTIVO

CHANGE POINT DETECTION CON RUPTURES



Utilizamos un algoritmo de detección de puntos de cambio que identifica un **cambio brusco en la tendencia** de la **probabilidad de incumplimiento en función del saldo**

Descripción	Valor
Modelo Utilizado	Binary Segmentation (Binseg) con L2
Número de puntos de cambio estimados	1
Punto de cambio (Saldo)	73



Podemos confirmar el resultado del algoritmo con los modelos anteriormente utilizados y de manera más clara con esta visualización de barras. Entre 0 y 100 empieza la escalada en % de morosidad: **73 es el umbral propuesto.**

¿Qué umbrales de saldo podrían indicar mayor riesgo de morosidad? **A PARTIR DE 73**