

# MINI-HACKATHON: DETECCIÓN DE FRAUDE

## Guía de Sprints Simplificada

Este hackathon NO busca la excelencia técnica. El objetivo es que **cuestionéis vuestras decisiones** y entendáis cómo diferentes departamentos trabajan juntos en un proyecto real.

### Contexto rápido

Sois un banco que quiere detectar fraude en compras online. Tenéis que decidir qué hacer con cada transacción:

APROBAR	Pasa sin problema
VERIFICAR	Pedir código SMS o confirmación en app (el cliente tarda 10-30 seg)
REVISAR	Un analista humano mira el caso (solo tenéis 80 análisis/día)
RECHAZAR	Se deniega el pago (¡cuidado! Si es un cliente bueno, se enfada)

#### Las restricciones del banco son:

- No podéis bloquear (rechazar) más del 2% de transacciones buenas
- Solo podéis revisar manualmente 80 casos al día
- El 5% de las transacciones del dataset son fraude real

## SPRINT 1: ENTENDER EL PROBLEMA

 25-30 minutos

Al final del sprint, cada equipo debe poder explicar el problema en 1 minuto.

### Equipo ML

#### TAREA 1: Elegir la métrica principal (10 min)

Debéis elegir UNA de estas tres métricas como la más importante para este proyecto:

Métrica	¿Qué mide?	¿Cuándo elegirla?
RECALL	% de fraudes que detectamos del total de fraudes reales	Si lo peor es que se nos escape fraude
PRECISION	% de alertas que son fraude real (no falsas alarmas)	Si lo peor es molestar a clientes buenos
F1	Equilibrio entre recall y precision	Si ambas cosas importan igual

#### TAREA 2: Identificar 3 señales de riesgo en los datos (10 min)

Mirad el CSV y pensad: ¿Qué 3 variables os parecen más útiles para detectar fraude? Apuntad por qué.

Variables disponibles: *importe\_eur, puntuacion\_riesgo\_ip, cambio\_dispositivo, antiguedad\_cuenta\_dias, transacciones\_ult\_24h, contracargos\_prev\_180d, hora\_dia, pais, categoria\_comercio*

### REUNIÓN con Riesgos (5 min)

Id a hablar con ellos y preguntadles:

1. "¿Qué es peor para el banco ahora: que se nos escape un fraude o que bloqueemos a un cliente bueno?"
2. "¿Hay alguna variable que os preocupe usar? (ej: ¿podemos usar el país?)"

→ Apuntad sus respuestas, las necesitaréis para defender vuestra métrica.

### Equipo Riesgos

#### TAREA 1: Identificar 3 cosas que pueden salir mal (10 min)

Pensad en qué podría ir mal si el sistema de fraude no está bien diseñado. Ejemplos para inspiraros:

- "Bloqueamos a todos los clientes de un país → nos acusan de discriminación"
- "Bloqueamos a viajeros que compran desde el extranjero → se enfadan y se van"

#### TAREA 2: Escribir 3 "reglas de oro" (10 min)

Escribid 3 reglas simples que el sistema SIEMPRE debe cumplir. Ejemplo:

- "Nunca rechazar solo por el país de origen"
- "Antes de rechazar a un cliente antiguo (>1 año), siempre pedir verificación primero"

### REUNIÓN con Producto (5 min)

Id a hablar con ellos y preguntadles:

3. "¿Cuánta fricción aguanta el cliente? ¿Pedir verificación al 5% de compras es demasiado?"
4. "Si os saturáis de casos manuales, ¿qué hacéis? ¿Aprobar todo? ¿Rechazar todo?"

## Equipo Operaciones

### TAREA 1: Dibujar el flujo de decisión (10 min) Podéis usar <https://app.diagrams.net/>

Dibujad un diagrama simple de qué pasa cuando llega una compra. Usad este esquema:

COMPRA → [Sistema evalúa] → ¿Riesgo? → Aprobar / Verificar / Revisar / Rechazar  
→ ¿Resultado?

### TAREA 2: Definir vuestros límites operativos (10 min)

Responded a estas preguntas con números:

- ¿Cuántas revisiones manuales podéis hacer al día?
- ¿Qué % máximo de compras podéis pedir verificación sin que los clientes se enfaden?

## REUNIÓN con Negocio (5 min)

Id a hablar con ellos y preguntadles:

5. "¿Qué preferís este trimestre: evitar más fraudes o no molestar a clientes buenos?"
6. "¿Cuál es el máximo de clientes buenos que podemos bloquear sin que sea un desastre?"

## Equipo Negocio

### TAREA 1: Definir 2 indicadores de éxito (10 min)

¿Cómo sabréis si el proyecto ha funcionado? Elegid 2 indicadores medibles. Ejemplos:

- "Reducir el fraude un X%"
- "No superar Z rechazos incorrectos al mes"

### TAREA 2: Establecer 1 restricción innegociable (10 min)

Definid UNA cosa que el sistema no puede hacer bajo ningún concepto. Ejemplo:

- "Los rechazos incorrectos no pueden superar el 2%"

## REUNIÓN con ML (5 min)

Id a hablar con ellos y preguntadles:

7. "¿Qué métrica vais a usar? ¿Por qué esa y no otra?"
8. "¿Qué variables del dataset os parecen más potentes para detectar fraude?"

## ENTREGABLE SPRINT 1 (todos los equipos juntos)

Rellenad esta ficha en un folio o pizarra:

- Métrica principal elegida: \_\_\_\_\_ (ML)
- 2 indicadores de éxito: \_\_\_\_\_ (Dirección)
- 1 restricción innegociable: \_\_\_\_\_ (Dirección)
- Capacidad revisión manual: \_\_\_\_\_ casos/día (Ops)
- 3 riesgos principales: \_\_\_\_\_ (Riesgos)

## SPRINT 2: DISEÑAR LA POLÍTICA DE DECISIÓN

⌚ 35-40 minutos

Decidir qué hacer con cada nivel de riesgo. Al final, tendréis una tabla de "si riesgo X → hacer Y".

### Equipo ML

#### TAREA 1: Crear un "score de riesgo" simple (15 min)

Sin entrenar ningún modelo, proponed una fórmula sencilla para calcular el riesgo. Ejemplo:

```
PUNTOS DE RIESGO = (puntuacion_riesgo_ip > 50 ? +2 : 0)
+ (cambio_dispositivo == 1 ? +2 : 0)
+ (antiguedad_cuenta_dias < 30 ? +1 : 0)
+ (transacciones_ult_24h > 5 ? +1 : 0)
```

Si PUNTOS >= 4 → ALTO | Si PUNTOS >= 2 → MEDIO | Si PUNTOS < 2 → BAJO

#### TAREA 2: Preparar la explicación (10 min)

Si un analista pregunta "¿por qué esta compra es sospechosa?", ¿qué 3 cosas le diríais? Pensad en variables fáciles de entender.

### REUNIÓN con Operaciones (5 min)

Presentadles vuestro sistema de puntos y preguntad:

9. "¿Os parece bien que MEDIO vaya a verificación y ALTO a revisión?"
10. "¿Preferís que el riesgo MEDIO vaya a verificación o directamente a revisión?"

### Equipo Riesgos

#### TAREA 1: Revisar la propuesta de ML (10 min)

Cuando el equipo de ML os presente su sistema de puntos, valorad si cumple vuestras "reglas de oro" del Sprint 1. Marcad:

- Esto está bien
- Esto nos preocupa porque...
- Esto no lo podemos permitir

#### TAREA 2: Definir 3 "casos protegidos" (10 min)

¿En qué casos NUNCA se debería rechazar directamente? Definid 3 situaciones donde siempre haya que verificar primero. Ejemplo:

- "Cliente con más de 2 años de antigüedad"

### REUNIÓN con Negocio (5 min)

Preguntadles para definir el equilibrio:

11. "¿Cuánto estás dispuestos a gastar en verificaciones extra para evitar 1€ de fraude?"
12. "Si hay que elegir: ¿menos fraude o menos clientes molestos?"

## Equipo Operaciones

### TAREA 1: Escribir la política final (15 min)

Con la info de ML y Riesgos, rellenad esta tabla:

Nivel de riesgo	Acción	¿Por qué?
BAJO	→	
MEDIO	→	
ALTO	→	
MUY ALTO	→	

### TAREA 2: Plan de saturación (10 min)

Si un día llegan más de 80 casos para revisión manual, ¿qué hacéis? Elegid una opción:

- Priorizar los de mayor importe y el resto a verificación
- Aprobar los casos más pequeños y revisar solo los grandes
- Otra opción: \_\_\_\_\_

## REUNIÓN con ML (5 min)

Preguntadles para ajustar la política:

13. "¿Qué señales son más fiables para no equivocarnos con clientes buenos?"
14. "¿Cómo sabríamos si estamos bloqueando demasiado?"

## Equipo Negocio

### TAREA 1: Elegir el equilibrio (15 min)

Debéis elegir UNA de estas dos estrategias:

ESTRATEGIA SEGURA	ESTRATEGIA CLIENTE
Priorizar detectar fraude aunque molestemos más a clientes (más verificaciones)	Priorizar no molestar a clientes aunque se cuelle algo de fraude

### TAREA 2: Poner números a los límites (10 min)

Definid con números concretos:

- % máximo de compras que pueden ir a verificación: \_\_\_\_\_%
- % máximo de compras que pueden ser rechazadas: \_\_\_\_\_%

## DILEMA con Riesgos (5 min)

Plantead este dilema a Riesgos:

*"Si para reducir el fraude a la mitad necesitamos pedir verificación al 10% de las compras... ¿es aceptable? ¿Qué problemas veis?"*

→ Escuchad sus objeciones y decidid juntos si merece la pena y por qué.

 **ENTREGABLE SPRINT 2 (todos los equipos juntos)**

1. Tabla de política: Riesgo → Acción (con justificación)
2. Límites numéricos: % verificación, % rechazo, casos revisión/día
3. Estrategia elegida: ¿Segura o Cliente? ¿Por qué?

## SPRINT 3: PLAN DE LANZAMIENTO

⌚ 25-30 minutos

Planificar cómo probamos el sistema, qué vigilamos, y qué hacemos si algo va mal.

### �� Equipo ML

#### TAREA 1: Elegir cómo probar el sistema (10 min)

¿Cómo lo lanzamos sin arriesgar demasiado? Elegid UNA opción:

A) Shadow mode	El sistema puntúa pero NO actúa. Comparamos sus predicciones con lo que habría pasado.
B) Piloto parcial	Activamos solo para un segmento (ej: compras >100€, o solo en España).

#### TAREA 2: Definir 4 cosas que vigilar (10 min)

¿Qué números miráis cada día para saber si el sistema va bien? Proponed 4:

- *Ejemplo: "% de fraudes detectados esta semana vs semana pasada"*

### 🤝 REUNIÓN con Negocio (5 min)

Preguntadles:

15. "¿Qué resultado os haría decir 'el proyecto ha sido un éxito' en 2 semanas?"
16. "¿Y en 2 meses? ¿Qué esperáis diferente?"

### 👑 Equipo Riesgos

#### TAREA 1: Checklist de cumplimiento (10 min)

¿Qué 5 cosas hay que tener preparadas antes de lanzar? Marcad con ✓:

- Registro de todas las decisiones (quién fue bloqueado y por qué)
- Proceso para que un cliente reclame si cree que fue injusto
- Revisión periódica de los umbrales (cada cuánto)
- \_\_\_\_\_
- \_\_\_\_\_

#### TAREA 2: Definir el "botón rojo" (10 min)

¿Qué condición o condiciones obliga a PARAR el sistema inmediatamente? Ejemplo:

- *"Si hay más de 10 quejas al día por bloqueos → revisar umbrales"*

### 🤝 REUNIÓN con Operaciones (5 min)

Preguntadles:

17. "¿Cómo gestionáis las reclamaciones de clientes bloqueados?"
18. "¿Qué mensaje mandáis al cliente cuando le pedís verificación?"

## Equipo Operaciones

### TAREA 1: Diseñar el día a día (10 min)

¿Cómo funciona la cola de revisión manual? Definid:

- ¿Por qué orden se revisan los casos? (¿mayor importe primero? ¿más antiguos?)
- ¿Cuánto tiempo máximo puede esperar un caso?
- ¿Qué pasa con los casos repetidos del mismo cliente?

### TAREA 2: Escribir el mensaje al cliente (10 min)

Escribid un mensaje corto y amable para cuando el cliente recibe una verificación:

*Ejemplo: "Para proteger tu cuenta, necesitamos verificar esta compra. Introduce el código que te hemos enviado al móvil."*

## REUNIÓN con ML (5 min)

Preguntadles:

19. "Si un cliente reclama, ¿qué explicación le damos? ¿Qué variables podemos mencionar?"
20. "¿Qué NO deberíamos decir para no revelar cómo funciona el sistema?"

## Equipo Negocio

### TAREA 1: Roadmap realista (10 min)

Definid qué se entrega en cada fase:

2 SEMANAS	MVP mínimo: ¿Qué entregamos SÍ O SÍ?
2 MESES	Versión mejorada: ¿Qué añadimos después?

### TAREA 2: Definir quién manda (5 min)

Responded:

- ¿Quién aprueba cambios en los umbrales?
- ¿Cada cuánto se revisa si el sistema funciona bien?

## VALIDACIÓN con Riesgos (5 min)

Pedid a Riesgos que os explique su "botón rojo":

*"¿Qué situación os obligaría a parar el sistema? ¿Estamos de acuerdo?"*

## ENTREGABLE SPRINT 3 (todos los equipos juntos)

1. Plan de pruebas: Shadow mode o piloto (cuál y por qué)
2. 4 métricas de monitorización
3. Condición del "botón rojo" (cuándo parar)
4. Roadmap: qué en 2 semanas, qué en 2 meses

