**CASO 4 – DIAGNÓSTICO INVISIBLE: Prevención de enfermedades crónicas con datos silenciosos**

Una ONG internacional llamada **VidaLatina** trabaja en proyectos de salud preventiva en regiones urbanas con alta prevalencia de enfermedades crónicas no transmisibles (diabetes, hipertensión, obesidad).

Tienen una base de datos **no clínica**: personas que aún no han sido diagnosticadas, pero sobre las cuales se tienen datos generales recolectados en ferias de salud, encuestas y pruebas de laboratorio básico (IMC, glucosa, frecuencia de ejercicio, consumo de azúcar, edad, etc.).

Tu equipo ha sido contratado para **detectar perfiles silenciosos en riesgo**, antes de que la enfermedad crónica aparezca formalmente. Pero hay una gran limitante: **no se puede usar directamente la variable ‘enfermedad\_cronica’ para decisiones clínicas**. Solo se permite usarla para análisis agregado poblacional. Además, te han pedido una propuesta para desplegar el modelo en campo (API o herramienta de captura).

**Desafíos**

1. ¿Cómo haces feature engineering para capturar patrones de riesgo sin tener historial clínico?
2. ¿Qué tipo de variable sería más útil: una dummy binaria, una interacción, o un binning transformado?
3. ¿Cómo haces que el modelo sea explicable para un voluntario sin formación médica?
4. ¿Qué pasa si tu modelo predice *falsos positivos de enfermedad*? ¿Generas pánico?
5. ¿Podrías proponer una forma de usar este modelo como herramienta digital en tiempo real?

**Reflexión ética**

¿Y si este modelo genera alarmismo en poblaciones vulnerables?  
¿Podría ser usado para excluir a pacientes de programas de seguros?  
¿Quién debe tener acceso al score de riesgo: el paciente, el médico, el gobierno?

**Estructura del Dataset**

| **Variable** | **Descripción** |
| --- | --- |
| edad | Edad en años |
| imc | Índice de Masa Corporal |
| fuma | 1 si fuma, 0 si no |
| ejercicio\_frecuencia | Categoría: Nunca, 1-2 veces, 3+ veces (por semana) |
| ingesta\_azucar | Cantidad estimada diaria de azúcar en gramos |
| enfermedad\_cronica | Variable objetivo (1 = diagnosticado, 0 = no diagnosticado aún) |