



<b>PROYECTO INTEGRADOR</b>
<b>Tema:</b> Sistema de Diagnóstico Médico con IA Reflexión sobre implicaciones Éticas del Caso “DermaDetect AI”

Profesora	GLADYS MARIA VILLEGAS RUGEL
Materia	PROYECTO INTEGRADOR EN INTELIGENCIA ARTIFICIAL
Alumnos	Bolaños Escandón María Fernanda Montaño Cárdenas Fernando Xavier
Fecha	07/10/2025

## Contenido

CONTEXTO DEL CASO .....	3
REFLEXIÓN SOBRE IMPLICACIONES ÉTICAS DEL CASO “DERMADETECT AI” .....	5
1. Propósito y Beneficios .....	5
2. Riesgos y Daños Potenciales.....	5
3. Fairness y Sesgo .....	6
4. Privacidad y Datos .....	7
5. Transparencia .....	7
6. Accountability.....	8
CONCLUSIÓN .....	8

## CONTEXTO DEL CASO

**Organización:** HealthAI Diagnostics (empresa ficticia basada en casos reales)

**Sistema:** DermaDetect AI

**Dominio:** Salud - Diagnóstico dermatológico

**Descripción del Sistema:** HealthAI Diagnostics ha desarrollado "DermaDetect AI", un sistema de diagnóstico asistido por IA para detectar cáncer de piel. El sistema:

- Analiza imágenes de lesiones cutáneas usando deep learning (CNN)
- Clasifica lesiones como: benignas, sospechosas o malignas • Genera reportes automáticos con nivel de confianza
- Recomienda acciones: monitoreo, biopsia o tratamiento inmediato
- Funciona como app móvil para pacientes y profesionales médicos

### Dataset de Entrenamiento

El modelo fue entrenado con:

- 100,000 imágenes de lesiones cutáneas
- Datos recopilados principalmente de hospitales en países desarrollados
- Composición racial: 85% pieles claras (tipos I-III Fitzpatrick), 15% pieles oscuras (tipos IV VI) • Imágenes capturadas con equipamiento médico profesional de alta calidad
- Etiquetas verificadas por dermatólogos certificados

### Resultados Reportados

La empresa reporta:

- Accuracy del 94% en pieles claras
- Accuracy del 78% en pieles oscuras
- Sensibilidad (recall): 92% en pieles claras, 71% en pieles oscuras
- Performance comparable a dermatólogos en pieles claras
- FDA no ha aprobado el sistema para uso clínico independiente

### Modelo de Negocio

- Versión gratuita: Pacientes toman fotos con smartphone, reciben clasificación automática
- Versión profesional: Dermatólogos pagan suscripción para integrar en su práctica
- Monetización: Datos de usuarios se utilizan para mejorar el modelo
- Términos de servicio: Usuarios ceden derechos sobre imágenes y datos

### **Implementación y Uso**

- 500,000 descargas de la app en 6 meses
- Disponible en 35 países, incluyendo regiones con acceso limitado a dermatólogos
- Muchos usuarios en países en desarrollo no tienen acceso a dermatólogos
- App no requiere verificación médica profesional para diagnósticos
- Marketing: "Detección de cáncer de piel al alcance de todos"

### **Incidentes Reportados**

Caso 1: Paciente en área rural de África con lesión maligna recibió clasificación "benigna" de la app. Esperó 8 meses antes de buscar atención médica. Cáncer se había expandido.

Caso 2: 15% de usuarios en encuesta admitieron NO buscar atención médica después de diagnóstico "benigno" de la app.

Caso 3: Estudio independiente encontró que el modelo tiene tasa de falsos negativos 3 veces mayor en pieles oscuras.

### **Respuesta de la Empresa**

HealthAI Diagnostics respondió: "DermaDetect es una herramienta de screening, no un reemplazo del diagnóstico profesional. Los usuarios aceptan términos que indican que deben consultar un médico. Estamos trabajando en mejorar el modelo para pieles oscuras en futuras versiones."

# REFLEXIÓN SOBRE IMPLICACIONES ÉTICAS DEL CASO “DERMADETECT AI”

## 1. Propósito y Beneficios

### ¿Cuál es el propósito principal?

El propósito principal de *DermaDetect AI* es facilitar la detección temprana de cáncer de piel mediante algoritmos de aprendizaje profundo que analizan imágenes dermatológicas. Busca ofrecer una herramienta accesible para pacientes y médicos, especialmente en zonas con escaso acceso a especialistas.

### ¿Quién se beneficia?

Los beneficiarios directos son los pacientes que podrían recibir una alerta temprana sobre posibles lesiones malignas, y los profesionales médicos que pueden usar la herramienta como apoyo diagnóstico. A nivel corporativo, también se beneficia HealthAI Diagnostics, que obtiene ingresos por suscripciones y mejora su base de datos para futuros desarrollos.

### ¿Los beneficios son equitativos?

No. Los beneficios son desiguales, pues el sistema muestra un rendimiento significativamente menor en personas con pieles oscuras. Esta brecha genera injusticia algorítmica, donde ciertos grupos raciales reciben diagnósticos menos precisos, vulnerando el principio de equidad en salud.

### ¿Hay propósitos alternativos?

Sí. Un propósito éticamente más sólido sería utilizar el sistema como herramienta complementaria en contextos clínicos, bajo supervisión médica y con modelos entrenados en datasets más diversos e inclusivos. También podría emplearse con fines educativos o de investigación, evitando su uso directo para diagnóstico sin validación humana.

## 2. Riesgos y Daños Potenciales

### ¿Qué puede salir mal?

El principal riesgo es el falso negativo en el diagnóstico de lesiones malignas, lo que puede retrasar el tratamiento y poner en riesgo la vida del paciente. También existe el riesgo de falsa confianza por parte del usuario, que podría interpretar el resultado como diagnóstico médico definitivo.

### **¿Quién podría ser perjudicado?**

Las personas con pieles oscuras son las más afectadas, ya que el modelo tiene una sensibilidad mucho menor en ese grupo. Además, pacientes sin acceso a dermatólogos podrían ser perjudicados al confiar exclusivamente en la app.

### **¿Cuáles son los peores escenarios?**

El peor escenario es la pérdida de vidas humanas por diagnósticos erróneos, junto con la erosión de la confianza pública en las tecnologías médicas basadas en IA. A nivel social, podría consolidarse una brecha digital sanitaria, donde los más vulnerables reciben peor atención.

### **¿Existen riesgos sistémicos?**

Sí. Si este tipo de herramientas se generaliza sin regulación, se podría instaurar un modelo médico desigual, donde la IA refuerza sesgos estructurales y desplaza la atención profesional en sectores marginados.

## **3. Fairness y Sesgo**

### **¿Hay sesgos en los datos?**

Sí. El dataset de entrenamiento contiene un 85% de imágenes de pieles claras, generando un sesgo demográfico que afecta la capacidad del modelo para generalizar a otras poblaciones.

### **¿El modelo discrimina grupos?**

De forma indirecta, sí. La menor precisión en pieles oscuras constituye una forma de discriminación algorítmica, ya que las decisiones del sistema son menos confiables para ciertos grupos raciales.

### **¿Perpetúa desigualdades existentes?**

Sí. Reproduce desigualdades históricas en investigación médica, donde las poblaciones no europeas o afrodescendientes están subrepresentadas en los estudios clínicos. Esto amplía la brecha sanitaria global.

### **¿Cómo se mitigan sesgos?**

La mitigación requiere reentrenar el modelo con datos representativos, incorporar técnicas de *fairness-aware learning* (machine learning variables sensibles, mejorar la equidad de los sistemas de IA) y realizar auditorías independientes. Además, se deben publicar métricas desagregadas por grupo demográfico para garantizar transparencia.

## 4. Privacidad y Datos

### ¿Qué datos se recopilan?

Se recopilan imágenes dermatológicas, metadatos del dispositivo, localización y posiblemente información personal del usuario. Estos datos se almacenan para mejorar el modelo y para fines comerciales.

### ¿Hay consentimiento informado?

El consentimiento es **insuficiente y ambiguo**. Los términos de servicio incluyen cesión total de derechos sobre imágenes, sin asegurar que el usuario comprenda las implicaciones. No hay evidencia de un consentimiento informado conforme a estándares éticos o de protección de datos (como el GDPR o la Ley de Protección de Datos Personales).

### ¿Cómo se protegen los datos?

No se especifican medidas claras de cifrado, anonimización ni políticas de retención. Esto plantea dudas sobre la seguridad y gobernanza de los datos médicos, que deberían tener protección reforzada.

### ¿Se minimizan los datos?

No. La empresa utiliza la información recolectada para entrenamiento continuo, lo que contradice el principio de minimización de datos, que exige recolectar solo lo estrictamente necesario para el propósito declarado.

## 5. Transparencia

### ¿Es explicable el modelo?

No completamente. Los modelos basados en redes neuronales convolucionales (CNN) son opacos por naturaleza. Los usuarios y médicos no pueden entender fácilmente por qué se clasifica una lesión como benigna o maligna.

### ¿Los usuarios entienden cómo funciona?

La mayoría no. La interfaz de la app no explica las limitaciones ni el grado de incertidumbre. El marketing **induce** a pensar que se trata de una herramienta diagnóstica confiable y directa.

### ¿Se comunican limitaciones?

Parcialmente. La empresa menciona en los términos legales que la app no reemplaza el diagnóstico médico, pero esta información no se presenta de forma visible o comprensible. Esto constituye una omisión ética relevante.

### **¿Hay auditoría posible?**

Actualmente no. No se han publicado los parámetros del modelo, los criterios de entrenamiento ni los datasets utilizados. La ausencia de auditoría independiente impide validar sus resultados y equidad.

## **6. Accountability**

### **¿Quién es responsable?**

La empresa HealthAI Diagnostics es responsable moral y legalmente del diseño, comercialización y mantenimiento del sistema. Sin embargo, su declaración pública busca trasladar la responsabilidad al usuario, lo cual es éticamente inaceptable.

### **¿Hay mecanismos de apelación?**

No existen mecanismos formales para que los usuarios reporten errores o impugnen diagnósticos. Esto vulnera el principio de rendición de cuentas establecido por la OCDE y la UNESCO para sistemas de IA de alto impacto.

### **¿Cómo se corrigen errores?**

La empresa indica que está “mejorando el modelo”, pero no hay un protocolo claro para notificar a usuarios afectados ni compensaciones ante daños. No existe un sistema de retroalimentación supervisado.

### **¿Existe supervisión humana?**

No en la versión gratuita. Esta ausencia constituye uno de los mayores fallos éticos del caso: la falta de intervención humana en decisiones críticas de salud, contraviniendo el principio de *Human-in-the-loop* recomendado por IEEE y la UE.

## **CONCLUSIÓN**

El caso *DermaDetect AI* ilustra cómo una innovación con alto potencial social puede transformarse en un problema ético y de justicia tecnológica cuando se prioriza la expansión comercial sobre la equidad, transparencia y responsabilidad. La solución requiere reentrenar el modelo con datos diversos, reforzar el consentimiento informado, implementar auditorías externas y garantizar supervisión humana permanente.

El desafío ético no es impedir la IA en salud, sino asegurar que cada avance tecnológico respete la dignidad, la igualdad y la vida humana.