

Aplicaciones de la analítica: Analítica en salud.

Laura Betancourt, Gilberto Gil, Aura Molina.

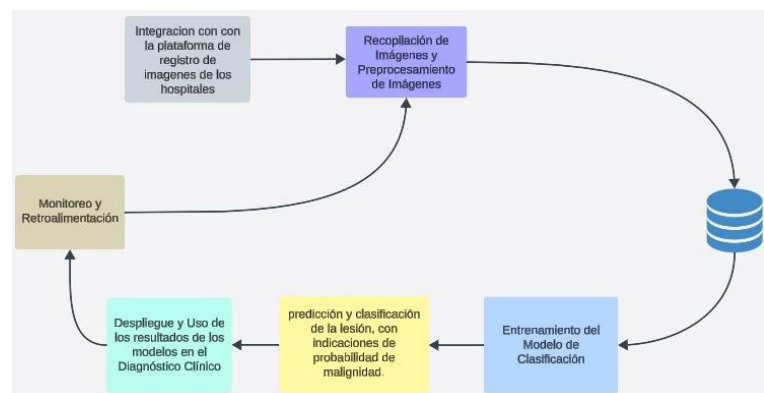
Caso de estudio: Cáncer de piel.

El cáncer de piel es una de las formas de cáncer más prevalentes, y su detección temprana es esencial para mejorar los resultados del tratamiento y la calidad de vida de los pacientes. Sin embargo, los métodos actuales de diagnóstico dependen en gran medida de la observación visual, lo que puede dar lugar a variaciones en el diagnóstico y tratamiento. Esto ha llevado a la necesidad de implementar herramientas analíticas avanzadas que permitan detectar de manera precisa si un lunar en la piel es maligno o benigno. Un modelo efectivo puede reducir la dependencia de métodos diagnósticos invasivos y mejorar la eficiencia en el diagnóstico, permitiendo a los profesionales de la salud concentrarse en casos de alto riesgo.

Problema analítico: desarrollar un modelo de clasificación de imágenes capaz de identificar con alta precisión si una lesión cutánea es maligna o benigna.

a. Diseño de solución propuesto.

El modelo clasificará las imágenes de lunares en “Lunar maligno”, “Lunar benigno” y “Revisión manual” siendo esta última la necesidad de que un experto revise manualmente la imagen, además se las imágenes que superen el 70% de probabilidad serán clasificadas en maligno, menos del 20% en benigno y el resto necesitaran de la revisión medica.



b. Análisis exploratorio.

La base de datos utilizada fue obtenida de Kaggle bajo el nombre '[Skin Cancer: Malignant vs Benign](#)',

Las imágenes tienen una resolución inicial de 224x224 píxeles, pero se redujeron a 100x100 píxeles para optimizar el procesamiento sin perder información relevante. Además, los valores de los píxeles, con un máximo de 255 por canal, fueron normalizados a un rango de 0 a 1 para mejorar la eficiencia en el entrenamiento del modelo.

c. Selección de algoritmos y técnicas de modelado.

Para abordar el problema de clasificación de lesiones cutáneas como malignas o benignas, se utilizaron varios algoritmos de modelado para evaluar cuál ofrecía el mejor desempeño en términos de precisión y capacidad predictiva. Los modelos seleccionados fueron:

- **XGBoost (Extreme Gradient Boosting)**

A pesar de que XGBoost no está diseñado para procesar imágenes directamente, se eligió este modelo en combinación con PCA (Análisis de Componentes Principales) para reducir la dimensionalidad de las imágenes. PCA extrae las características más relevantes y las transforma en un vector compacto, lo que permite a XGBoost trabajar con datos más manejables y menos ruidosos. Esta combinación aprovecha la eficiencia de XGBoost para clasificación, manteniendo un bajo costo computacional y mejorando el rendimiento del modelo sin necesidad de redes neuronales profundas.

- **Red neuronal estándar**

La elección de una red neuronal estándar se justifica por su capacidad para aprender patrones complejos y no lineales directamente desde los datos. Las redes neuronales pueden manejar relaciones intrincadas entre los píxeles de una imagen y extraer características relevantes sin necesidad de un procesamiento previo exhaustivo. Además, gracias a su estructura flexible, son eficaces para tareas de clasificación, como la detección de cáncer de piel, donde la variabilidad visual es alta y se requieren modelos capaces de generalizar adecuadamente.

- **Red neuronal convolucional CNN**

La elección de una CNN se justifica debido a su capacidad para aprender y reconocer patrones espaciales en imágenes de manera automática. Las CNN son especialmente efectivas para tareas de clasificación de imágenes como la detección de cáncer de piel, ya que pueden capturar características locales y globales de las imágenes, como texturas y formas, que son esenciales para distinguir entre lesiones benignas y malignas. Esta capacidad de aprendizaje jerárquico permite a las CNN ofrecer una alta precisión y eficiencia en el análisis de datos visuales complejos.

d. Comparación de modelos y métricas.

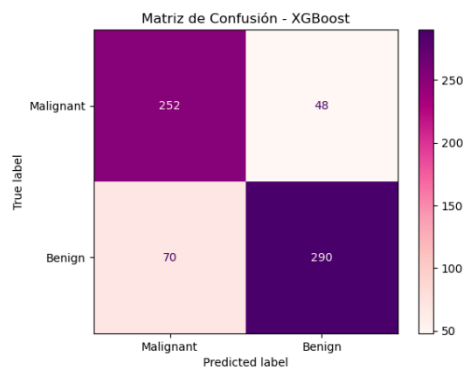
En este caso se utiliza Recall como métrica principal en el modelo de clasificación debido a su capacidad para capturar la mayor cantidad posible de casos positivos verdaderos, lo que permite minimizar los falsos negativos. Los falsos negativos representan un riesgo significativo en la problemática, ya que corresponden a instancias en las que el modelo no identifica una condición de enfermedad cuando en realidad está presente. Esto podría llevar a que ciertos pacientes no reciban un tratamiento oportuno, lo cual pone en riesgo su salud. A continuación, se muestran los resúmenes de las métricas, tanto para los datos de train y test de los tres modelos entrenados.

-----TRAIN NEURAL NETWORK-----					-----TRAIN XGBOOST-----					-----TRAIN CNN-----						
83/83	2s 16ms/step					precision	recall	f1-score	support		precision	recall	f1-score	support		
	0	0.86	0.84	0.85	1440	0	1.00	1.00	1.00	1440	0	0.82	0.84	0.83	1440	
	1	0.81	0.84	0.83	1197	1	1.00	1.00	1.00	1197	1	0.80	0.78	0.79	1197	
	accuracy				2637	accuracy				2637	accuracy				2637	
	macro avg				0.84	macro avg				1.00	macro avg				0.81	2637
	weighted avg				0.84	weighted avg				1.00	weighted avg				0.81	2637
AUC - Train Neural Network: 0.8398678989139515					AUC - Train XGBoost: 1.0					Test AUC: 0.8096321821212769						
-----TEST NEURAL NETWORK-----					-----TEST XGBOOST-----					-----TEST CNN-----						
21/21	0s 10ms/step					precision	recall	f1-score	support		precision	recall	f1-score	support		
	0	0.81	0.83	0.82	360	0	0.84	0.81	0.83	360	0	0.83	0.83	0.83	360	
	1	0.79	0.77	0.78	300	1	0.78	0.82	0.80	300	1	0.79	0.80	0.80	300	
	accuracy				660	accuracy				660	accuracy				660	
	macro avg				0.80	macro avg				0.81	macro avg				0.81	660
	weighted avg				0.80	weighted avg				0.81	weighted avg				0.81	660
AUC - Test Neural Network: 0.7972222222222222					AUC - Test XGBoost: 0.8138888888888889											

e. Afinamiento de hiperparámetros.

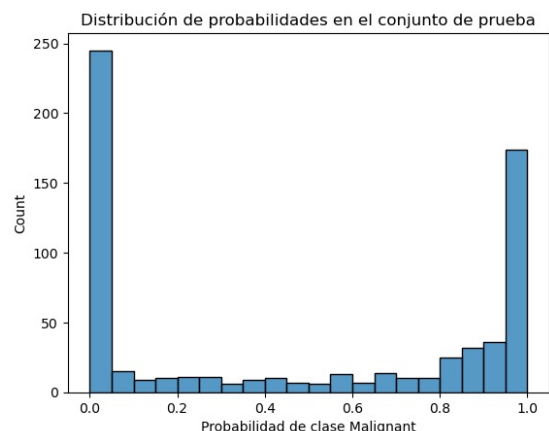
Para XGBoost, se utilizó el método RandomizedSearchCV para explorar diferentes combinaciones de hiperparámetros, como el número de árboles, la profundidad máxima de los árboles, la tasa de aprendizaje, y otros parámetros clave como el porcentaje de muestras utilizadas en cada árbol. El objetivo era optimizar el modelo para mejorar su capacidad de generalización y su rendimiento en la clasificación de lesiones cutáneas. Tras afinar estos hiper parámetros, el modelo presenta mejoras en las métricas de interés, a continuación, el resumen de estas.

-----TRAIN XGBOOST-----				
	precision	recall	f1-score	support
0	1.00	1.00	1.00	1440
1	1.00	1.00	1.00	1197
accuracy			1.00	2637
macro avg	1.00	1.00	1.00	2637
weighted avg	1.00	1.00	1.00	2637
AUC - Train XGBoost: 1.0				
-----TEST XGBOOST-----				
	precision	recall	f1-score	support
0	0.86	0.81	0.83	360
1	0.78	0.84	0.81	300
accuracy			0.82	660
macro avg	0.82	0.82	0.82	660
weighted avg	0.82	0.82	0.82	660



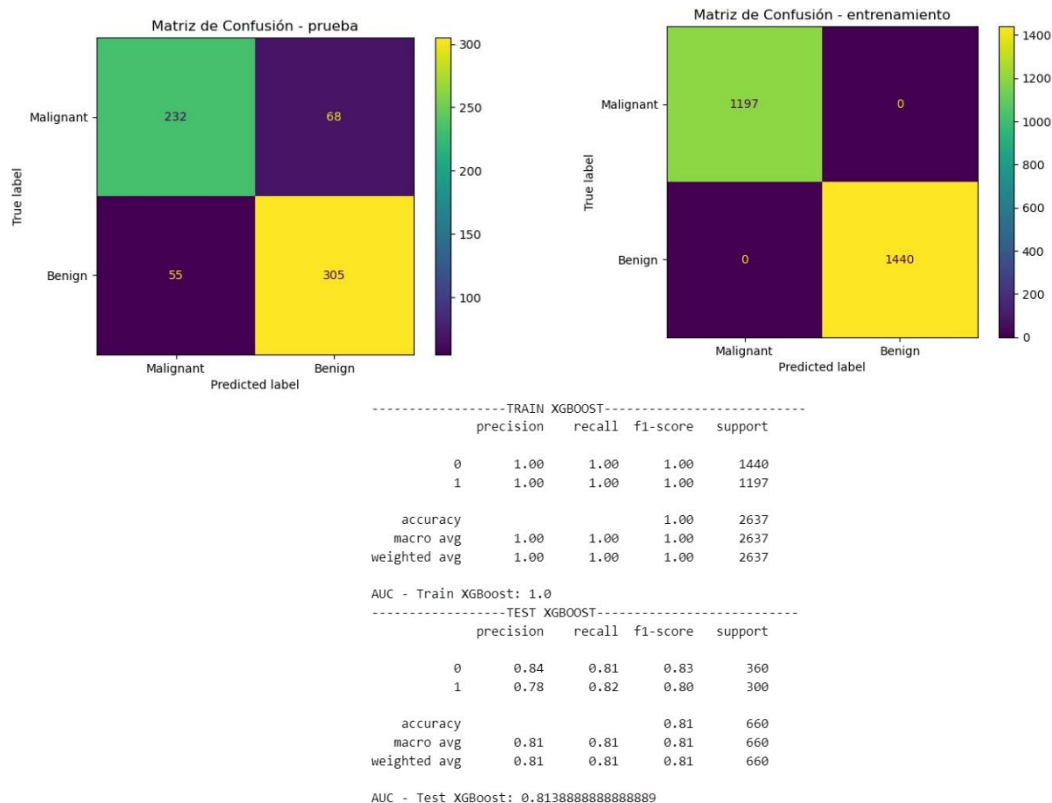
f. Evaluación y análisis del modelo.

Al observar la distribución de las Proporciones, se observa que estas se concentran en los valores extremos. Como consecuencia, se toma la decisión de adoptar un enfoque riguroso en las decisiones, en el que se clasifique con mucha seguridad. Se decide clasificar como "lunar maligno" únicamente si la probabilidad es igual o mayor al 98%, y como "lunar benigno" si la probabilidad es igual o menor al 2%.



Con este enfoque, el modelo clasifica con alta seguridad como que tiene un lunar benigno el 42.88%, y como que sí tienen un lunar maligno el 43.94%. El restante 13.18% de las personas evaluadas se selecciona para una revisión manual por parte de especialistas, en función del orden de probabilidad de tener un lunar maligno que puede ser un indicador de cáncer de piel, dado que es crítico tanto evitar informar erróneamente la presencia de cáncer como omitir su detección.

Dicho lo anterior, esta solución permite que los médicos encargados de revisar las imágenes de los lunares solo tengan que revisar aproximadamente el 13.18% de estas, disminuyendo en gran medida la ocupación de los profesionales, el tiempo de diagnóstico para el paciente y también los altos costos para la organización de salud.



Análisis de los que sí tienen Lunares malignos

De las personas que tienen Lunares malignos, el modelo es capaz de predecir correctamente el 82%. De todas las predicciones obtenidas con el modelo para la clase positiva, solo son erróneas el 20%. Esto nos permite estar razonablemente seguros de las clasificaciones que se les dan a las personas con cáncer de piel. El 18% de error en el recall podría corresponder a pacientes con probabilidades cercanas al 98%, lo que significa que están en el límite de ser clasificados como positivos, pero no alcanzan el umbral necesario para la clasificación positiva en el contexto definido.

Análisis de los que tienen lunares benignos

De las personas que tienen lunares benignos, el modelo es capaz de predecir correctamente el 81%. De todas las predicciones obtenidas para esta clase, solo el 19% es erróneo. Esto proporciona un alto grado de seguridad en las clasificaciones de personas sin cáncer de piel. El 19% de error en el recall para esta clase (falsos negativos) podría corresponder a pacientes con probabilidades cercanas al 2%. Estas probabilidades son bajas, pero no lo suficiente para clasificarlos como negativos con total seguridad, debido a que no alcanzan el umbral para una clasificación clara de "sin lunares benignos".

g. Despliegue del modelo.

El despliegue permite analizar imágenes de lunares mediante un modelo de aprendizaje automático. Primero, las imágenes se cargan, se convierten al formato RGB y se redimensionan, tras lo cual el modelo clasifica cada lunar como 'benigno', 'Maligno' o 'requiere revisión'. A partir de esta clasificación, se genera un mensaje personalizado para cada caso: si es "benigno", se recomienda mantener chequeos regulares, si es 'revisión', se sugiere consultar a un especialista y finalmente si es 'Maligno' se genera una alerta de consulta al médico especialista de inmediato. Además, se muestra un porcentaje de confianza en cada diagnóstico, lo que permite entender la seguridad del modelo en sus predicciones. Finalmente, los resultados se guardan en un archivo .xlsx para su análisis y uso en decisiones médicas posteriores.

Paciente	Clase	Mensaje	Probabilidad
10	lunar Maligno	Lamentamos informarle que hemos detectado ciertas anomalías en la evaluación de su lunar que sugieren la posibilidad de malignidad. Sabemos que esta noticia puede ser abrumadora, pero es crucial que actúe con prontitud. Le recomendamos que se comunique de inmediato con su médico para realizar una evaluación detallada y discutir las opciones de tratamiento. Estamos aquí para apoyarle en cada paso del proceso. Su bienestar es nuestra prioridad.	99.41%
100	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de malignidad. Puede sentirse tranquilo/a, ya que su salud está en buen estado. Sin embargo, le recomendamos continuar con chequeos regulares para mantener su bienestar.	0.92%
1000	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.02%
1001	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.0%
1002	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	2.08%
1004	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.87%
1005	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.21%
1006	lunar Maligno	Lamentamos informarle que hemos detectado ciertas anomalías en la evaluación de su luna	97.97%
1007	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.07%
1008	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.41%
1009	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.07%
101	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.17%
1010	lunar Benigno	¡¡Buenas noticias! La evaluación de su lunar indica que es benigno y no presenta signos de n	0.0%