

Clasificación de cáncer de piel utilizando aprendizaje profundo Skin cancer classification using deep learning

L. R. Jaime-Calva ^a, H. Castillejos-Fernández ^{a,*}, A. Franco-Arcega ^a, P. Miranda-Romagnoli ^b, O. Pérez-Cortés ^a

^aÁrea Académica de Computación y Electrónica, Universidad Autónoma del Estado de Hidalgo, 42184, Pachuca, Hidalgo, México.

^bÁrea Académica de Matemáticas y Física, Universidad Autónoma del Estado de Hidalgo, 42184, Pachuca, Hidalgo, México.

Resumen

El cáncer de piel es una enfermedad que afecta a las personas con tono de piel oscuro o claro. Por otra parte, cada vez más personas tienden a emplear camas de bronceado o permanecer periodos prolongados de tiempo ante los rayos del sol, provocando que esta enfermedad sea más frecuente. Como complemento en el diagnóstico de esta enfermedad existe la inteligencia artificial, la cual, permite emplear algoritmos de clasificación como árboles de decisión, bosques aleatorios, regresión logística, entre otros; además, del uso de algoritmos de aprendizaje profundo como las redes neuronales convolucionales, ayudando a realizar un pre-diagnóstico de cáncer de piel. En este artículo se explicará el desarrollo de un método, en el cual, utilizando la base de datos de imágenes dermatológicas publicada por el Internacional Skin Imaging Collaboration (ISIC) (Collaboration, 2022) se considera un conjunto de imágenes las cuales ya han sido caracterizadas por especialistas y que se encuentran en un grupo de benignas y malignas.

Palabras Clave: Aprendizaje Automático, Aprendizaje Profundo, Clasificación, Redes Neuronales Convolucionales

Abstract

Skin cancer is a disease that affects people with dark or light skin tone. On the other hand, more and more people tend to use tanning beds or spend prolonged periods of time in the sun's rays, causing this disease to be more frequent. As a complement in the diagnosis of this disease there is artificial intelligence, which allows the use of classification algorithms such as decision trees, random forest, logistic regression, among others; In addition, the use of deep learning algorithms such as convolutional neural networks, helping to make a pre-diagnosis of skin cancer. This article will explain the development of a method, in which, using the database of dermatological images published by the International Skin Imaging Collaboration (ISIC) (Collaboration, 2022), a set of images is considered which have already been characterized by specialists and found in a group of benign and malignant.

Keywords: Machine Learning, Deep Learning, Classification, Convolutional Neural Networks

1. Introducción

El cáncer de piel es el crecimiento descontrolado de células anómalas de la piel dando lugar a tumores malignos. El 5% de los cánceres de piel son melanomas (crecimiento descontrolado de los melanocitos), y en el otro 95% se engloban el carcinoma basocelular (afectación de las células basales) o el espinocelular (multiplicación de las células escamosas) Pacientes (2022).

Según Foundation (2022), sobreviene cuando el daño en el ADN de las células de la piel desencadena mutaciones o defectos genéticos que hacen que las células de la piel se multipliquen rápidamente dando lugar a tumores malignos. Algunos tipos de cáncer de piel son Queratosis Actínica, Carcinoma Ba-

socelular, Nevos Displásicos, Melanoma, Carcinoma de Células Escamosas y Carcinoma de Células de Merkel. El cáncer de piel puede afectar a cualquier persona; sin embargo, muchos pacientes y médicos creen que las personas con piel oscura son inmunes a la enfermedad, esto no es verdad y esa creencia puede generar peligrosas demoras en el diagnóstico. Por esto, aunque los Caucásicos son más propensos a desarrollar cáncer de piel, las personas con piel oscura que lo desarrollan tienen una mayor posibilidad de morir por esta causa. Se sabe que la piel tiene tres capas principales: Epidermis, dermis y grasa subcutánea. Un tipo de células de la epidermis son los melanocitos, los cuales producen melanina, el pigmento que da el color a la

*Autor para correspondencia: heydy_catillejos@uaeh.edu.mx

Correo electrónico: ja454709@uaeh.edu.mx (Luis Rey Jaime-Calva), heydy_catillejos@uaeh.edu.mx (Heydy Castillejos-Fernández), afranco@uaeh.edu.mx (Anil Franco-Arcega), pmiranda@uaeh.edu.mx (Pedro Miranda-Romagnoli), obed_perez@uaeh.edu.mx (Obed Pérez-Cortés)

Historial del manuscrito: recibido el 18/04/2022, última versión-revisada recibida el 03/06/2022, aceptado el 02/08/2022, publicado el 31/08/2022. **DOI:** <https://doi.org/10.29057/icbi.v10iEspecial3.9029>



piel, a los ojos y al pelo. La melanina ayuda a proteger la piel de los rayos ultravioleta (RUV), una conocida causa de cáncer de piel. Las personas que producen menos cantidad de melanina son más vulnerables a los rayos del sol o a las cámaras bronceadoras.

De acuerdo a (Clinic, 2022) algunos factores que pueden aumentar el riesgo de sufrir cáncer de piel son:

- Piel clara. Independientemente del color de la piel, toda persona puede contraer cáncer de piel. Sin embargo, tener menos pigmento (melanina) en la piel proporciona menos protección contra la radiación UV perjudicial.
- Antecedentes de quemaduras por sol. Haber tenido una o más ampollas de quemaduras por el sol de niño o adolescente aumenta el riesgo de desarrollar cáncer de piel de adulto.
- Exposición excesiva al sol. Cualquier persona que pase un tiempo considerable al sol puede desarrollar cáncer de piel, especialmente si la piel no está protegida con protección solar o ropa.
- Climas soleados o en grandes altitudes. Las personas que viven en climas cálidos y soleados están expuestas a mayor luz solar que las personas que viven en climas más fríos.
- Lunares. Las personas que tienen muchos lunares o lunares atípicos denominados nevos displásicos se encuentran en mayor riesgo de padecer cáncer de piel.

Las tasas más altas de incidencia del melanoma maligno cutáneo se registran en Australia y Nueva Zelanda, seguidos de América del Norte y Europa del Norte Roldan (2017).

Para el año 2022, de acuerdo a Society (2022) los cálculos en los Estados Unidos son:

- Se diagnosticarán aproximadamente 99,780 nuevos casos de melanoma (alrededor de 57,180 en hombres y 42,600 en mujeres)
- Se calcula que morirán aproximadamente 7,650 personas (5,080 hombres y 2,570 mujeres) a causa de melanoma.

Desde 2001, el cáncer de piel es el segundo más frecuente en México. Además de que es responsable del 80 % de las muertes por cáncer de piel en México Roldan (2017).

La International Skin Imaging Collaboration (ISIC) es una asociación académica e industrial diseñada para facilitar la aplicación de imágenes digitales de la piel para ayudar a reducir la mortalidad por melanoma. El principal objetivo clínico de ISIC es apoyar los esfuerzos para reducir las muertes relacionadas con el melanoma y las biopsias innecesarias al mejorar la precisión y la eficiencia de la detección temprana del melanoma. ISIC crea recursos para las comunidades de dermatología e informática, incluido un gran archivo de acceso público de código abierto y en expansión de imágenes de la piel. Este archivo sirve como recurso público de imágenes para la enseñanza, la investigación y el desarrollo y prueba de algoritmos de inteligencia artificial de diagnóstico. El Archivo ISIC es una plataforma de código abierto con imágenes disponibles públicamente de lesiones cutáneas bajo licencias Creative Commons. Las imágenes

están asociadas con diagnósticos reales y otros metadatos clínicos. Collaboration (2022)

En la figura 1 y 2 se presentan algunos ejemplos de imágenes con lesiones benignas y malignas respectivamente.

2. Aprendizaje automático

El aprendizaje automático investiga cómo las computadoras pueden aprender (o mejorar su rendimiento) en función de los datos. Un área de investigación principal es que los programas de computadora aprendan automáticamente a reconocer patrones complejos y tomen decisiones inteligentes basadas en datos Han and Kamber (2006).

Los sistemas de aprendizaje automático se pueden clasificar según la cantidad y el tipo de supervisión que reciben durante el entrenamiento. Hay cuatro categorías principales: aprendizaje supervisado, aprendizaje no supervisado, aprendizaje semi-supervisado y aprendizaje por refuerzo Géron (2019).

En este trabajo se emplea el aprendizaje supervisado, en el cual, los datos de entrenamiento que alimenta al algoritmo incluyen las soluciones deseadas, llamadas etiquetas. Una tarea típica de aprendizaje supervisado es la clasificación Géron (2019).

Los algoritmos de clasificación usan el aprendizaje supervisado junto con varias clases de datos: el objetivo es entrenar un modelo que pueda hacer predicciones categóricas. Los clasificadores binarios funcionan con conjuntos de datos que tienen dos clases, mientras que los clasificadores multiclase (a veces llamados clasificadores multinomiales) distinguen más de dos clases Campesanto (2020).

Dentro de los algoritmos de clasificación se encuentran los árboles de decisión, la regresión logística, k-vecinos más cercanos, bosques aleatorios, entre otros.

2.1. Árboles de decisión

Los árboles de decisión son otro tipo de algoritmo de clasificación que involucra una estructura similar a un árbol. En un árbol genérico, la ubicación de un punto de datos está determinada por una lógica condicional simple. Un árbol de decisión crea un clasificador en forma de árbol. Este se compone de nodos de decisión, donde se realizan pruebas sobre atributos específicos; y nodos hoja, que indican el valor del atributo objetivo. Para clasificar una nueva muestra, se comienza en la raíz del árbol y se navega por los nodos hasta llegar a una hoja Campesanto (2020).

2.2. Bosques aleatorios

Los bosques aleatorios son una generalización de los árboles de decisión: este algoritmo de clasificación involucra múltiples árboles. Si los datos implican hacer una predicción numérica, se calcula el promedio de las predicciones de los árboles. Si los datos involucran una predicción categórica, se determina el modo de las predicciones de los árboles Campesanto (2020).



Figura 1: Ejemplos de imágenes de clase benigna.

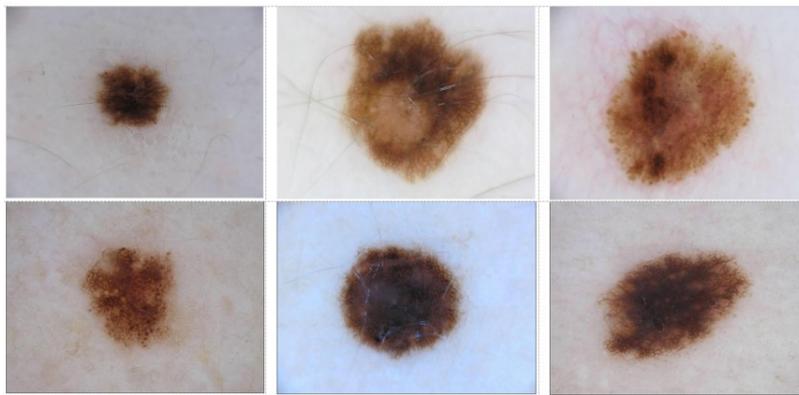


Figura 2: Ejemplos de imágenes de clase maligna.

2.3. Regresión logística

La regresión logística es una clasificación y un modelo lineal con una salida binaria. La regresión logística funciona con múltiples variables independientes e involucra una función sigmoidea para calcular probabilidades. La regresión logística es esencialmente el resultado de aplicar la función de activación sigmoidea a la regresión lineal para realizar la clasificación binaria. La regresión logística puede ser binomial (solo dos resultados para una variable dependiente), multinomial (tres o más resultados para una variable dependiente) u ordinal (las variables dependientes están ordenadas) Campesanto (2020).

2.4. K Vecinos más cercanos

El algoritmo k-vecinos más cercanos (en inglés, k-nearest neighbors, abreviado k-NN) es un algoritmo de clasificación. En el que, los puntos de datos que están cerca uno del otro se clasifican como pertenecientes a la misma clase. Cuando se introduce un nuevo punto, se agrega a la clase de la mayoría de su vecino más cercano. El algoritmo k-NN es esencialmente una técnica heurística y no con fundamentos matemáticos complejos y, sin embargo, sigue siendo un algoritmo efectivo y útil Campesanto (2020).

2.5. Aprendizaje profundo

El aprendizaje profundo es un subcampo específico del aprendizaje automático: una nueva forma de aprender representaciones a partir de datos que pone énfasis en aprender capas

sucesivas de representaciones cada vez más significativas. En el aprendizaje profundo, estas representaciones en capas se aprenden (casi siempre) a través de modelos llamados redes neuronales, estructurados en capas literales apiladas una encima de la otra. El término red neuronal es una referencia a la neurobiología, pero aunque algunos de los conceptos centrales del aprendizaje profundo se desarrollaron en parte inspirándose en la comprensión del cerebro, los modelos de aprendizaje profundo no son modelos del cerebro. El aprendizaje profundo es un marco matemático para aprender representaciones a partir de datos Chollet (2018).

2.5.1. Red neuronal convolucional

Una red neuronal convolucional (en inglés, convolutional neural network, abreviado CNN) es un perceptrón multicapa diseñado específicamente para reconocer formas bidimensionales con un alto grado de invariancia a la traducción, escala, sesgo y otras formas de distorsión. Esta difícil tarea se aprende de manera supervisada por medio de una red cuya estructura incluye las siguientes formas de restricciones Chollet (2018):

- 1. Extracción de características. Cada neurona toma sus entradas sinápticas de un campo receptivo local en la capa anterior, obligándola así a extraer características locales. Una vez que se ha extraído una característica, su ubicación exacta se vuelve menos importante, siempre que se conserve aproximadamente su posición relativa a otras características.

- 2. Mapeo de funciones. Cada capa computacional de la red se compone de múltiples mapas de características, y cada mapa de características tiene la forma de un plano dentro del cual las neuronas individuales están restringidas para compartir el mismo conjunto de pesos sinápticos. Esta segunda forma de restricción estructural tiene los siguientes efectos beneficiosos:

- Cambio de invariancia, forzado en la operación de un mapa de características mediante el uso de convolución con un núcleo de pequeño tamaño, seguido de una función sigmoidea;
- Reducción del número de parámetros libres, lograda mediante el uso de pesos compartidos.

- 3. Submuestreo. A cada capa convolucional le sigue una capa computacional que realiza promedios y submuestreos locales, por lo que se reduce la resolución del mapa de características. Esta operación tiene el efecto de reducir la sensibilidad de la salida del mapa de características a los cambios y otras formas de distorsión.

Una característica definitoria importante de las redes neuronales convolucionales es una operación, que se conoce como convolución. Una operación de convolución es una operación de producto escalar entre un conjunto de pesos con estructura de cuadrícula y entradas similares con estructura de cuadrícula extraídas de diferentes localidades espaciales en el volumen de entrada. En las redes neuronales convolucionales, los estados de cada capa se organizan de acuerdo con una estructura de cuadrícula espacial. Estas relaciones espaciales se heredan de una capa a la siguiente porque cada valor de característica se basa en una pequeña región espacial local en la capa anterior. Es importante mantener estas relaciones espaciales entre las celdas de la cuadrícula, porque la operación de convolución y la transformación a la siguiente capa dependen de manera crítica de estas relaciones. Cada capa en la red convolucional es una estructura de cuadrícula tridimensional, que tiene una altura, un ancho y una profundidad Aggarwal (2018).

Los tres tipos de capas que comúnmente están presentes en una red neuronal convolucional son convolución, agrupación y ReLU. En la red neuronal convolucional, los parámetros se organizan en conjuntos de unidades estructurales tridimensionales, conocidas como filtros o núcleos.

3. Criterios de evaluación

En la mayoría de las aplicaciones prácticas, se dividen los datos disponibles en tres conjuntos. El primer conjunto se usa para entrenar el algoritmo y el segundo se usa como conjunto de validación para la selección del modelo. Después de seleccionar el mejor modelo, se prueba el rendimiento del predictor de salida con el tercer conjunto, que a menudo se denomina conjunto de prueba. El número obtenido se utiliza como estimador del error verdadero del predictor aprendido S. Shalev-Shwartz (2014).

3.1. Exactitud

La exactitud mide una fracción de las predicciones del clasificador que son correctas. Se calcula con la ecuación 1:

$$Exactitud = \frac{VP + VN}{VP + VN + FP + FN}, \quad (1)$$

donde, VP = Verdadero Positivo que es el número de predicciones positivas correctas. VN = Verdadero Negativo que es el número de predicciones negativas correctas. FP = Falso Positivo que es el número de predicciones positivas falsas. FN = Falso Negativo que es el número de predicciones negativas falsas.

3.2. Precisión

La precisión se puede considerar como una medida de exactitud que se calcula con la ecuación 2:

$$Precision = \frac{VP}{VP + FP}. \quad (2)$$

3.3. Sensibilidad

La sensibilidad es la fracción de las instancias verdaderamente positivas que reconoce el clasificador. Una puntuación de sensibilidad de uno indica que el clasificador no hizo predicciones negativas falsas. Se calcula con la ecuación 3:

$$Sensibilidad = \frac{VP}{VP + FN}. \quad (3)$$

3.4. Validación Cruzada

En la validación cruzada de k veces, los datos iniciales se dividen aleatoriamente en k subconjuntos mutuamente excluyentes o "pliegues", D_1, D_2, \dots, D_k cada uno de aproximadamente el mismo tamaño. El entrenamiento y las pruebas se realizan k veces Han and Kamber (2006).

El propósito de la validación cruzada es probar un modelo con conjuntos de prueba que no se superponen, lo cual se realiza de la siguiente manera:

- Paso 1) Dividir los datos en k subconjuntos de igual tamaño.
- Paso 2) Seleccione un subconjunto para probar y otros para entrenar.
- Paso 3) Se repite el paso 2 para otros subconjuntos $k - 1$.

Este proceso se denomina validación cruzada de k veces, y la estimación de error general es el promedio de la estimación de error. Un método estándar para la evaluación implica una validación cruzada de diez veces. Extensos experimentos han demostrado que 10 subconjuntos es la mejor opción para obtener una estimación precisa Campesanto (2020).

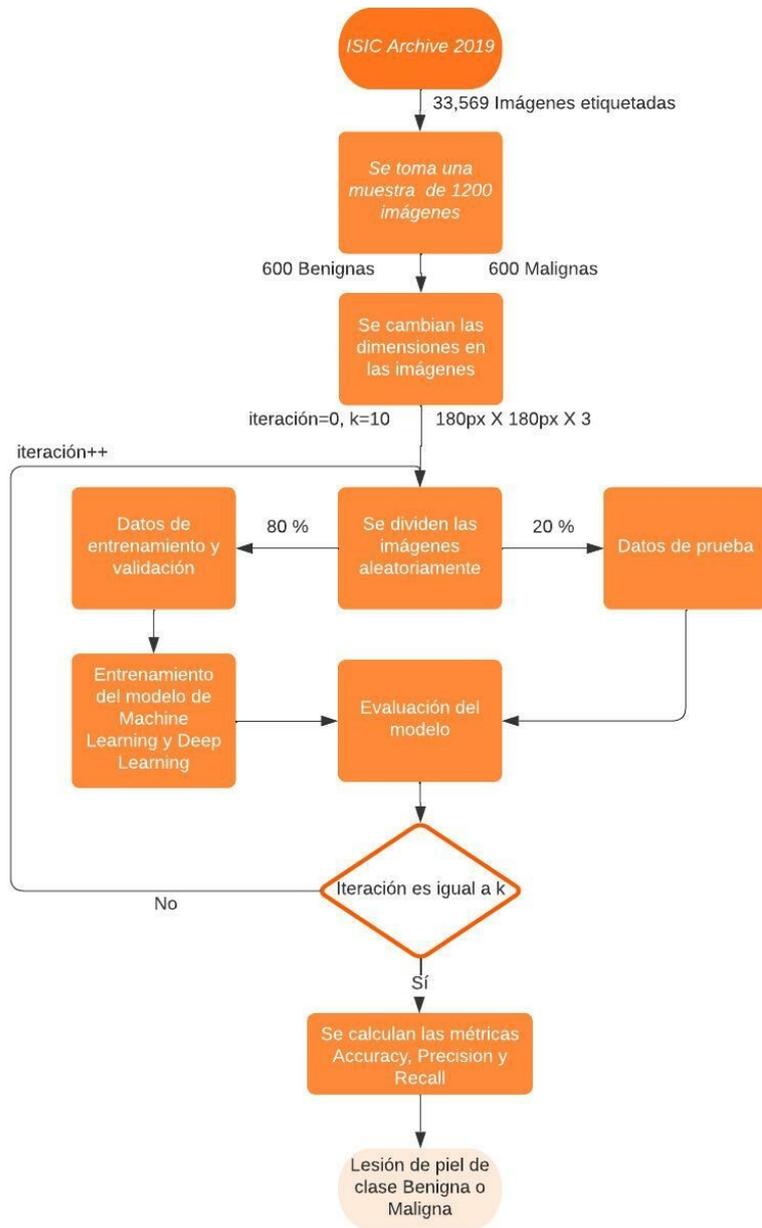


Figura 3: Diagrama del modelo propuesto.

4. Experimentos y resultados

Utilizando el archivo ISIC del año 2019, el cual consta de un total de 33569 imágenes divididas en 25331 para el conjunto de entrenamiento y 8238 para el conjunto de prueba, se tomo una muestra de 1200 imágenes, de las cuales 600 corresponden a lesiones benignas y 600 a lesiones malignas. La muestra de imágenes fue dividida en 80 % para entrenamiento y 20 % para prueba. Posteriormente, se creó un modelo de red neuronal convolucional constituido por 3 capas convolucionales, 3 capas de submuestreo y 2 capas densas, utilizando como función de activación la unidad lineal rectificadora (en inglés, rectified linear unit, abreviado ReLU).

Por otro lado también se utilizaron los algoritmos de clasificación árboles de decisión, regresión logística, bosques aleatorios y k vecinos más cercanos.

En ambos casos, tanto para los algoritmos de clasificación como para la red neuronal convolucional se utilizó el método de validación cruzada con 10 pliegues.

El diagrama del modelo propuesto se muestra en la figura 3. Por otra parte, los resultados obtenidos para cada clasificador se pueden consultar en la tabla 1.

Tabla 1: Resultados obtenidos

Métrica / Método	Exactitud	Precisión	Sensibilidad
Árboles de decisión	67.70 %	63.49 %	67.03 %
Regresión logística	72.05 %	73.56 %	71.11 %
Bosques aleatorios	76.00 %	74.73 %	81.60 %
k-NN	72.97 %	73.80 %	69.66 %
CNN	92.10 %	92.30 %	91.50 %

5. Discusión de resultados

Al comparar los resultados de la tabla 1, en donde se muestran los valores de los criterios de evaluación obtenidos para cada clasificador empleado, se puede ver que el modelo de red neuronal convolucional fue el mejor de todos obteniendo los siguientes resultados: exactitud de 92.20 %, precisión de 92.30 % y sensibilidad de 91.50 %. Lo cual indica un buen rendimiento para clasificar si la lesión de piel presente en una imagen es de clase benigna o maligna. Cabe destacar que este modelo de aprendizaje profundo ya tiene implícito un paso para la extracción de características, por lo cual, aprende las características de las imágenes y crea nuevas características automáticamente, lo cual no está presente en los demás algoritmos vistos en la literatura y se debe realizar como un paso extra. Esto nos permite ver las capacidades que tiene el modelo de red neuronal

convolucional sobre los algoritmos de aprendizaje automático clásicos.

6. Conclusión

En este artículo se empleó un modelo de aprendizaje profundo basado en redes neuronales convolucionales para clasificar imágenes de lesiones de piel y determinar si la lesión es de clase benigna o maligna. De igual forma, se comparó el rendimiento de este modelo con otros clasificadores encontrados en la literatura.

Por otra parte, para evaluar el rendimiento de los algoritmos de clasificación utilizados, se emplearon las métricas de exactitud, precisión y sensibilidad, además de utilizar el método de validación cruzada de k veces para obtener mejores estimaciones.

Como trabajo futuro se espera realizar nuevos experimentos incrementando la cantidad de imágenes utilizadas para los conjuntos de entrenamiento y prueba, así como, se van a modificar los parámetros utilizados para los algoritmos de clasificación como por ejemplo el número de capas convolucionales o las funciones de activación en el caso del algoritmo de red neuronal convolucional; o el número de árboles para el algoritmo de bosques aleatorios.

Referencias

- Aggarwal, C. C., editor (2018). *Neural Networks and Deep Learning*. Springer, Yorktown Heights, NY, USA.
- Campesanto, O., editor (2020). *Artificial Intelligence, Machine Learning and Deep Learning*. Mercury Learning and Information, Dulles, VA 20166.
- Chollet, F., editor (2018). *Deep Learning with Python*. Manning, Shelter Island, NY 11964.
- Clinic, M. (2022). Cancer de piel. [Web; accedido el 07-04-2022].
- Collaboration, S. I. (2022). About isic. [Web; accedido el 07-04-2022].
- Foundation, S. C. (2022). Cancer de piel en personas de piel oscura. [Web; accedido el 07-04-2022].
- Géron, A., editor (2019). *Hands-on Machine Learning with Scikit-Learn, Keras and Tensorflow*. O'Reilly, Gravenstein Highway North, Sebastopol, CA 95472.
- Han, J. and Kamber, M., editors (2006). *Data Mining Concepts and Techniques*. Elsevier, 500 Sansome Street, Suite 400, San Francisco, CA 94111.
- Pacientes, R. (2022). Cancer de piel. [Web; accedido el 07-04-2022].
- Roldan, R. (2017). Unam dirección general de comunicación social. [Web; accedido el 07-04-2022].
- S. Shalev-Shwartz, S. B.-D., editor (2014). *Understanding Machine Learning*. Cambridge University Press, 32 Avenue of the Americas, New York, NY 10013-2473, USA.
- Society, A. C. (2022). Estadísticas importantes sobre cáncer de piel tipo melanoma. [Web; accedido el 07-04-2022].