



Sistema de aprendizaje automático para la detección de cáncer de piel



SISTEMA DE APRENDIZAJE AUTOMÁTICO PARA LA DETECCIÓN DE CÁNCER DE PIEL

ELABORADO POR:
STEFANIA SCARLET HERNÁNDEZ QUIROZ
ALEJANDRO FIDEL LARA MIRANDA



UNIVERSIDAD POPULAR DEL CESAR
FACULTAD DE INGENIERÍAS Y TECNOLÓGICAS
INGENIERÍA DE SISTEMAS
VALLEDUPAR, CESAR
SEPTIEMBRE DE 2024

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

SISTEMA DE APRENDIZAJE AUTOMÁTICO PARA LA DETECCIÓN DE CÁNCER DE PIEL

**ELABORADO POR:
STEFANIA SCARLET HERNÁNDEZ QUIROZ
ALEJANDRO FIDEL LARA MIRANDA**

**DIRECTOR DEL PROYECTO:
ING. BRAULIO BARRIOS ZUÑIGA**

**UNIVERSIDAD POPULAR DEL CESAR
FACULTAD DE INGENIERÍAS Y TECNOLÓGICAS
INGENIERÍA DE SISTEMAS
VALLEDUPAR, CESAR
SEPTIEMBRE DE 2024**

TABLA DE CONTENIDO

	Pág.
RESUMEN	6
ABSTRACT	7
INTRODUCCIÓN	8
SECCIÓN I: DESCRIPCIÓN GENERAL	9
1.1.- TÍTULO DEL PROYECTO	9
1.2.- DIRECCIÓN DE EJECUCIÓN DEL PROYECTO	9
1.3.- LAPSO DE EJECUCIÓN DEL PROYECTO	9
1.4.- ORGANISMO Y SECCIÓN RESPONSABLE	3
1.5.- INFORMACION DE CONTACTO DE LOS ESTUDIANTES	3
1.6.- LÍNEA, SUBLÍNEA Y GRUPO DE INVESTIGACIÓN AL QUE SE SUSCRIBE EL PROYECTO	3
SECCIÓN II. DESCRIPCIÓN SITUACIONAL	3
2.1.- IDENTIFICACIÓN DEL PROBLEMA	3
2.2.- JUSTIFICACIÓN DE PROYECTO	5
2.3.- OBJETIVOS DEL PROYECTO	6
2.4 MARCO REFERENCIAL	6
2.4.1 MARCO LEGAL	8
2.4.2 GLOSARIO TÉCNICO	9
2.5 MARCO METODOLÓGICO	10
2.5.1 POBLACIÓN Y MUESTRA	11
2.5.2 INSTRUMENTOS Y TÉCNICAS PARA LA RECOLECCIÓN DE LA INFORMACIÓN	12
2.6 RESULTADOS ESPERADOS	12
2.7 PRESUPUESTO	14
SECCIÓN III: DESARROLLO CIENTÍFICO – TECNOLÓGICO	16
3.1 – DESARROLLO DE LAS FASES DE LA METOLOGÍA PROPUESTA	16
3.1.1 Introducción al desarrollo metodológico	16
3.1.1.1 Definición de roles	16
3.1.1.2 Definición de responsabilidades	18
3.1.1.3 Implementación de metodología Kanban	20
3.1.1.3.1 Creación del proyecto en Jira	21
3.1.1.3.2 Definición de columnas del tablero Kanban	21
3.1.1.3.3 Creación de historias de usuario y tareas	22
3.1.1.3.4 Asignación de tareas a miembros del equipo	22

3.1.1.3.5 Visualización del flujo de trabajo	22
3.1.1.3.6 Ejecución de tareas	22
3.1.1.3.7 Revisión y aprobación	22
3.1.1.3.8 Comentarios y colaboración	23
3.1.1.3.9 Seguimiento del progreso	23
3.1.2 Etapa de investigación	24
3.1.2.1 Fuentes de investigación	24
3.1.2.2 Desarrollo de la investigación preliminar	25
3.1.2.2.1 Bases de datos públicas de imágenes médicas	25
3.1.2.2.2 Entrevistas con ingenieros de sistemas	25
3.1.2.2.3 Revisión de literatura científica	25
3.1.2.3 Recolección de información	26
3.1.2.3.1 Diseño del formulario de preguntas	26
3.1.2.3.2 Implementación del formulario	27
3.1.2.4 Resultados obtenidos	29
3.1.2.4.1 Herramientas de análisis	29
3.1.2.5 Análisis de resultados	33
3.1.3 Preparación y planificación	36
3.1.3.1 Requerimientos funcionales	36
3.1.3.2 Requerimientos no funcionales	37
3.1.3.3 Historias de usuario	39
3.1.4 Base de datos	40
3.1.4.1 Descripción general	40
3.1.4.2 Datos del conjunto de imágenes	41
3.1.4.3 Clasificación de imágenes	41
3.1.4.4 Estructura de almacenamiento	41
3.1.4.5 Descripción de la base de datos ISIC	42
3.1.4.6 Origen y acceso a la base de datos	42
3.1.4.7 Relevancia para el proyecto	42
3.1.5 Desarrollo del aplicativo	43
3.1.5.1 Diseño de interfaces de usuario	43
3.1.5.2 Desarrollo del backend	47
3.1.5.2.1 Integraciones	48
3.1.5.3 Desarrollo del frontend	49
3.2 – ANÁLISIS DE RESULTADOS Y DISCUSIÓN	50
3.2.1 Relación con el objetivo general	50
3.2.2 Evaluación respecto a los objetivos específicos	50
3.2.3 Discusión sobre la metodología aplicada	51

3.3 – CONCLUSIONES	51
---------------------------	-----------

3.4 – RECOMENDACIONES	52
------------------------------	-----------

TABLA DE ILUSTRACIONES

Ilustración 1 Cronograma de actividades	14
Ilustración 2: Ejemplo de tablero Kanban	21
Ilustración 3: Cronograma de actividades Jira	23
Ilustración 4: Encuesta para la recolección de información	28
Ilustración 5: Gráfica de resultados del formulario de recolección de datos 1	29
Ilustración 6: Gráfica de resultados del formulario de recolección de datos 2	29
Ilustración 7: Gráfica de resultados del formulario de recolección de datos 3	30
Ilustración 8: Gráfica de resultados del formulario de recolección de datos 4	30
Ilustración 9: Gráfica de resultados del formulario de recolección de datos 5	30
Ilustración 10: Gráfica de resultados del formulario de recolección de datos 6	31
Ilustración 11: Gráfica de resultados del formulario de recolección de datos 7	31
Ilustración 12: Gráfica de resultados del formulario de recolección de datos 8	31
Ilustración 13: Gráfica de resultados del formulario de recolección de datos 9	32
Ilustración 14: Gráfica de resultados del formulario de recolección de datos 10	32
Ilustración 15: Gráfica de resultados del formulario de recolección de datos 11	32
Ilustración 16: Gráfica de resultados del formulario de recolección de datos 12	33
Ilustración 17 Interfaz de navegación inicial	43
Ilustración 18 Menú desplegable con las opciones a tomar	44
Ilustración 19 Interfaz de entrenamiento	44
Ilustración 20 Interfaz de entrenamiento	45
Ilustración 21 Interfaz de evaluación	45
Ilustración 22 Interfaz de selección de evaluación	46
Ilustración 23 Interfaz de resultado positivo de la predicción	46
Ilustración 24 Interfaz de resultado negativo de la predicción	47

INDICES DE TABLAS

Tabla 1 Presupuesto de recursos físicos	14
Tabla 2: Presupuesto de recursos tecnológicos	15
Tabla 3: Presupuesto de recursos humanos	15
Tabla 4: Roles del equipo	17
Tabla 5: Listado de actividades	18
Tabla 6: Formulario de recolección de información	26
Tabla 7 Preguntas de la encuesta que se realizo	33
Tabla 8 Requerimientos funcionales	36
Tabla 9 Requerimientos no funcionales	38
Tabla 10 Historia de usuario: Visualización de información sobre tipos de cáncer de piel	39
Tabla 11 Historia de usuario: Selección de algoritmo para entrenamiento	39
Tabla 12 Historia de usuario: Entrenamiento del modelo seleccionado	40
Tabla 13 Historia de usuario: Evaluación de imágenes de cáncer de piel	40
Tabla 14 Cantidad de imágenes por cada tipo de cáncer	41

RESUMEN

Este proyecto propone el diseño y desarrollo de un sistema integral de aprendizaje automático para la detección temprana de cáncer de piel, utilizando un enfoque basado en redes neuronales profundas. Con el objetivo de mejorar la precisión y la accesibilidad en el diagnóstico de esta enfermedad, el sistema está orientado tanto a profesionales de la salud como a usuarios en general que buscan una herramienta confiable para la evaluación inicial de lesiones cutáneas.

El proyecto parte de la selección de un banco de datos de imágenes para el diagnóstico de cáncer de piel, específicamente utilizando el conjunto de datos ISIC disponible en Kaggle, reconocido por su calidad y diversidad de casos. Posteriormente, se llevará a cabo el preprocesamiento y procesamiento de estas imágenes, garantizando que los datos estén optimizados para la entrada en las diferentes arquitecturas de redes neuronales.

Se evaluarán múltiples modelos de redes neuronales supervisadas para identificar la arquitectura más eficaz en términos de precisión diagnóstica. Tras seleccionar el modelo más adecuado, se procederá a su entrenamiento utilizando el conjunto de datos preprocesado. El sistema incluirá módulos especializados para la carga y procesamiento de imágenes, entrenamiento del modelo, simulación y generación de resultados.

Finalmente, se implementará un sistema web intuitivo y accesible que permitirá a los usuarios cargar imágenes de lesiones cutáneas, procesarlas mediante el modelo entrenado y obtener resultados rápidos y precisos sobre la probabilidad de malignidad. Este sistema no solo busca mejorar la detección temprana y reducir la carga de trabajo en los profesionales de la salud, sino también democratizar el acceso a herramientas avanzadas de diagnóstico, contribuyendo así a la mejora de la salud pública.

ABSTRACT

This project proposes the design and development of an integrated machine learning system for the early detection of skin cancer, using an approach based on deep neural networks. With the aim of improving accuracy and accessibility in the diagnosis of this disease, the system is geared toward both healthcare professionals and the general public who seek a reliable tool for the initial evaluation of skin lesions.

The project begins with the selection of a dataset of images for the diagnosis of skin cancer, specifically utilizing the ISIC dataset available on Kaggle, which is recognized for its quality and diversity of cases. Subsequently, the preprocessing and processing of these images will be carried out, ensuring that the data is optimized for input into the various neural network architectures.

Multiple supervised neural network models will be evaluated to identify the most effective architecture in terms of diagnostic accuracy. After selecting the most suitable model, it will be trained using the preprocessed dataset. The system will include specialized modules for image upload and processing, model training, simulation, and result generation.

Finally, an intuitive and accessible web system will be implemented that will allow users to upload images of skin lesions, process them through the trained model, and obtain quick and accurate results regarding the probability of malignancy. This system not only seeks to improve early detection and reduce the workload of healthcare professionals but also to democratize access to advanced diagnostic tools, thereby contributing to the improvement of public health.

INTRODUCCIÓN

El proyecto titulado "Sistema de Aprendizaje Automático para la Detección Temprana de Cáncer de Piel" se presenta como una solución tecnológica avanzada, destinada a revolucionar la manera en que se detecta y diagnostica el cáncer de piel. Situado en un contexto donde la detección temprana es crucial para el éxito del tratamiento, este proyecto aborda desafíos críticos relacionados con la precisión diagnóstica y la accesibilidad a herramientas médicas avanzadas. El sistema propuesto no solo busca mejorar la experiencia de los profesionales de la salud en la identificación de lesiones cutáneas malignas, sino también empoderar a los usuarios comunes con una herramienta confiable y de fácil acceso para la evaluación inicial de su salud dermatológica.

El corazón de este proyecto radica en el desarrollo de un sistema de aprendizaje automático basado en redes neuronales profundas, optimizado para analizar imágenes dermatológicas con alta precisión. A través de un proceso riguroso que incluye la selección de un banco de datos extenso y variado, el preprocesamiento de imágenes, y la evaluación y entrenamiento de múltiples arquitecturas de redes neuronales, el proyecto garantiza que el sistema sea robusto y efectivo. Este enfoque asegura que el modelo final, implementado en una plataforma web accesible, pueda proporcionar diagnósticos rápidos y precisos, accesibles para un amplio espectro de usuarios, desde médicos especialistas hasta individuos preocupados por su salud.

La relevancia del proyecto se sustenta en la identificación clara de una necesidad crítica: la falta de herramientas diagnósticas accesibles y precisas para la detección temprana del cáncer de piel. Esta carencia no solo limita la capacidad de respuesta ante casos potencialmente graves, sino que también crea una barrera significativa para la prevención efectiva de la enfermedad. En respuesta a esta situación, el sistema propuesto busca democratizar el acceso a tecnologías de diagnóstico avanzadas, permitiendo a los usuarios tomar decisiones informadas sobre su salud y contribuyendo a la reducción de la carga sobre los sistemas de salud pública.

Desde una perspectiva metodológica, el proyecto se apoya en un enfoque sistemático que combina técnicas avanzadas de procesamiento de imágenes con modelos de aprendizaje profundo para optimizar la precisión diagnóstica. La implementación del sistema web se realiza con un enfoque en la usabilidad, asegurando que la interfaz sea intuitiva y que los resultados sean fáciles de interpretar para todos los usuarios. Además, el proyecto incluye módulos especializados que permiten la carga, procesamiento, simulación y generación de

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

resultados, asegurando una experiencia de usuario fluida y eficiente.

SECCIÓN I: DESCRIPCIÓN GENERAL

1.1.- TÍTULO DEL PROYECTO

Sistema de aprendizaje automático para la detección de cáncer de piel

1.2.- DIRECCIÓN DE EJECUCIÓN DEL PROYECTO

Universidad Popular del Cesar sede Sabanas Diagonal 21 No. 29-56 Sabanas del Valle -
Valledupar- Cesar, teléfono (605)-5850573 ext 1111, pqrs@unicesar.edu.co.

1.3.- LAPSO DE EJECUCIÓN DEL PROYECTO

El proyecto sistema de aprendizaje automático para la detección de cáncer de piel será desarrollado en un lapso de cinco meses (5 meses).

Fase de definición y planificación (1 de enero – 31 de enero):

- Realizar la revisión bibliográfica y contextualización del problema.
- Definir los objetivos y alcance del proyecto.
- Construir el diseño preliminar del sistema y selección de herramientas y tecnologías.

Fase de recopilación de datos y preprocesamiento (1 de febrero – 29 de febrero):

- Identificar fuentes de datos relevantes para la detección de cáncer de piel.
- Recopilar datos y realizar preprocesamiento para garantizar la calidad.
- Desarrollar un conjunto de datos de entrenamiento.

Fase de desarrollo del modelo de aprendizaje automático (1 de marzo – 31 de marzo):

- Implementar algoritmos de aprendizaje automático para la detección.
- Entrenar el modelo utilizando el conjunto de datos.
- Evaluar inicialmente el rendimiento del modelo y realizar ajustes necesarios.

Fase de integración y pruebas del sistema (1 de abril – 30 de abril):

- Integrar el modelo en un sistema completo de detección.
- Desarrollar la interfaz de usuario y la experiencia del usuario.
- Realizar pruebas de funcionamiento y resolver posibles problemas.

Fase de optimización, documentación y presentación (1 de mayo – 31 de mayo)

- Optimizar el sistema y realizar ajustes finales.
- Documentar detalladamente el proyecto y los resultados obtenidos.
- Preparar la presentación final y entregar el proyecto.

1.4.- ORGANISMO Y SECCIÓN RESPONSABLE

Este proyecto se llevará a cabo de manera independiente, utilizando una base de datos pública disponible en Kaggle (referencia: [<https://www.kaggle.com/code/sharanharsoor/skin-cancer-detection/input>]). Dada la naturaleza abierta del proyecto y la ausencia de una entidad específica, el sistema se centrará en la creación de un prototipo que será evaluado exhaustivamente utilizando la base de datos mencionada. Estos repositorios son reconocidos por la calidad y fiabilidad de la información que albergan, y su uso se ajusta a los lineamientos éticos y académicos establecidos por la institución.

1.5.- INFORMACION DE CONTACTO DE LOS ESTUDIANTES

Nombre	Apellido	Cédula	Teléfono	Correo
Stefania Scarlet	Hernández Quiroz	100337792 4	302276467 9	sscarlethernandez@unices ar.edu.co
Alejandro Fidel	Lara Miranda	100428287 4	300579151 8	afidellara@unicesar.edu.co

1.6.- LÍNEA, SUBLÍNEA Y GRUPO DE INVESTIGACIÓN AL QUE SE SUSCRIBE EL PROYECTO

Línea de investigación: Transformación digital

Sublínea de investigación: Big Data y Analytics

Grupo de investigación: GISICO

SECCIÓN II. DESCRIPCIÓN SITUACIONAL

2.1.- IDENTIFICACIÓN DEL PROBLEMA

En la actualidad, la detección de cáncer de piel ha experimentado avances significativos gracias a la implementación de herramientas de inteligencia artificial, este enfoque se ha convertido en una base para el diagnóstico temprano y preciso de diversas enfermedades dermatológicas, incluido el cáncer de piel. En Colombia, donde las tasas de mortalidad por cáncer continúan siendo una preocupación de salud pública, la aplicación de estas tecnologías se vuelve crucial. La capacidad de realizar diagnósticos precisos mediante el análisis de imágenes dermatológicas con algoritmos de aprendizaje automático puede marcar la diferencia en la atención médica, permitiendo una intervención oportuna y mejorando las tasas de supervivencia. [1]

La situación del cáncer de piel en Colombia plantea una seria preocupación en salud pública, como lo indican las estadísticas del Instituto Nacional de Cancerología. Con 712 nuevos casos registrados en 2017, el cáncer de piel lidera las patologías tumorales malignas en la institución, destacándose el carcinoma basocelular como el más frecuente con un 52,7% de incidencia. Aunque los datos son limitados, se estima que la incidencia de melanoma maligno es significativa, alcanzando 4,6 por 100.000 en mujeres y 4,4 por 100.000 en hombres. Estas cifras subrayan la necesidad de estrategias efectivas de prevención y detección temprana para hacer frente a este creciente problema de salud en el país. [2]

El Instituto Nacional del Cáncer [3] ha informado que el melanoma es uno de los tipos de cáncer más graves. Esto sucede cuando demasiadas luces ultravioletas inciden en la piel, con el tiempo, se daña el ADN de las células de la piel, causa cáncer y lleva muchos años. La detección del cáncer de piel se basa en gran medida en la evaluación visual de lesiones cutáneas por parte de los médicos, lo que puede llevar a errores de diagnóstico debido a la variabilidad en la interpretación de las imágenes de las lesiones.

En este momento las instituciones prestadoras de servicio de tratamiento de cáncer de piel no cuentan con un sistema que implemente inteligencia artificial para la temprana y oportuna

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

detección de cáncer de piel. La falta de herramientas de apoyo a la toma de decisiones, como

modelos de pronóstico basados en datos clínicos y características de las lesiones, dificulta la estimación precisa de la progresión de la enfermedad. Así mismo el no contar con un sistema efectivo de seguimiento de pacientes y un pronóstico preciso dificulta la implementación oportuna de tratamientos y terapias adecuadas.

Si en la ciudad de Valledupar no se implementa una solución efectiva para este problema, es probable que siga enfrentando déficit para la detección oportuna de cáncer en la piel, causando así que se sigan presentando errores de diagnósticos debido a fallas humanas en la interpretación de las imágenes dermatoscopias.

Para enfrentar esta problemática de manera efectiva, es esencial la creación de una aplicación dedicada que incorpore tecnologías de aprendizaje automático para la detección de cáncer de piel. Esta herramienta no solo tiene el potencial de superar las limitaciones actuales, sino que también se posiciona como una solución integral para las necesidades de numerosos pacientes. Al mejorar la precisión en la detección, la aplicación busca minimizar tanto los falsos positivos como los falsos negativos, estableciendo así un sistema altamente confiable que puede marcar la diferencia en la lucha contra el cáncer de piel.

FORMULACION DEL PROBLEMA:

¿Cómo a través de la implementación de un sistema de aprendizaje automático para la detección de cáncer de piel se puede mejorar el pronóstico temprano y preciso?

2.2.- JUSTIFICACIÓN DE PROYECTO

Desde un punto de vista teórico, en el ámbito de la detección del cáncer de piel, la integración de la Inteligencia Artificial (IA) ha marcado un hito significativo. Este avance se ha materializado en investigaciones que exploran el potencial de los algoritmos de IA para diagnosticar melanomas. Mediante la combinación de técnicas de procesamiento de imágenes y aprendizaje automático, se ha logrado una alta capacidad de discriminación entre lesiones malignas y benignas, lo que sugiere un prometedor horizonte de aplicaciones clínicas [4]. Este enfoque de implementar la IA para la detección de cáncer de piel a través de imágenes planea desarrollar una solución que arroje resultados con mayor rapidez, eficaz y seguro.

Metodológicamente, este proyecto supone una oportunidad para aplicar y perfeccionar técnicas como: aprendizaje automático, procesamiento de imágenes dermatoscópicas, extracción de características, entrenamiento del modelo de aprendizaje, gestión de bases de datos y diseño de interfaces de usuario. La creación de una plataforma web funcional y eficiente requerirá un enfoque meticuloso en el diseño y desarrollo, lo que contribuirá al avance de habilidades y conocimientos en el campo de la Transformación digital y Big Data y Analytics.

Desde una perspectiva práctica, la implementación de una herramienta de inteligencia artificial en Valledupar es de suma importancia para mejorar la precisión y eficiencia en el diagnóstico del cáncer de piel. Esta herramienta permitirá la carga de imágenes dermatoscópicas, su procesamiento y análisis mediante algoritmos avanzados, y la generación de resultados precisos y rápidos. Esta solución integral no solo facilitará la detección temprana y permitirá tratamientos más oportunos, sino que también tendrá un impacto positivo en las tasas de supervivencia, abordando directamente las necesidades de salud de la comunidad. Además, la creación de una plataforma web funcional y eficiente superará las limitaciones actuales en la atención médica, proporcionando un análisis más objetivo y consistente de las imágenes dermatológicas. Esto reducirá la posibilidad de errores de diagnóstico y establecerá un sistema altamente confiable para la lucha contra el cáncer de piel. [5]

2.3.- OBJETIVOS DEL PROYECTO

- **Objetivo General:**

Desarrollar un sistema para la detección de cáncer de piel utilizando técnicas de aprendizaje automático.

- **Objetivos Específicos:**

1. Seleccionar el banco de datos de imágenes para el diagnóstico de cáncer de piel.
2. Realizar el preprocesamiento y procesamiento de las imágenes del banco de datos seleccionado
3. Evaluar las diferentes arquitecturas o modelos de redes neuronales de aprendizaje supervisado con el banco de datos seleccionado.
4. Realizar el entrenamiento de la red neuronal seleccionada utilizando el banco de datos de imágenes.
5. Desarrollar los módulos para la carga, procesamiento de imágenes, entrenamiento, simulación y generación de resultados.
6. Implementar el sistema web para la detección de cáncer de piel.

2.4 MARCO REFERENCIAL

1. **Aplicación de la inteligencia artificial en la clasificación automática de imágenes médicas:** Durante los últimos 15 años, un número creciente de técnicas referentes a imágenes digitales ha sido introducido en la práctica médica. Uno de los métodos de aprendizaje automático más utilizado en la actualidad son las redes neuronales. La detección temprana del melanoma es esencial para un tratamiento exitoso. Algunos signos de advertencia del melanoma incluyen un cambio en el tamaño, forma o color de un lunar existente, la aparición de un nuevo lunar que parece diferente de los demás, picazón, sangrado o ulceración de una lesión cutánea, y la propagación del pigmento desde el borde de una lesión hacia la piel circundante. En el presente trabajo se presentan resultados de la clasificación de imágenes médicas, específicamente con presencia de cáncer de piel, utilizando aprendizaje automático e inteligencia artificial. Se utilizaron dos bases de datos

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

para llevar a cabo la clasificación de las imágenes, ambas bases de datos contienen imágenes dermatológicas y se encuentran disponibles de manera gratuita en internet, lo cual facilita la comparación de resultados entre distintos autores. Para los experimentos se utilizó una red

neuronal la cual permite obtener resultados competitivos [6].

2. **Inteligencia artificial y aprendizaje automático en medicina:** El aprendizaje automático es una poderosa rama de la Inteligencia Artificial que se ha utilizado con éxito en distintas industrias. En los últimos años con la creciente disponibilidad de información clínica almacenada electrónicamente el campo médico se ha convertido en un ambiente ideal para el desarrollo y aplicación de estas nuevas tecnologías. El aprendizaje automático posee el potencial de mejorar los sistemas de salud, mediante el análisis de millones de datos clínicos se logran crear modelos pronósticos, de tamizaje y diagnósticos. Sin embargo, a pesar de ser evidente que el uso de métodos algorítmicos puede mejorar la calidad de los sistemas de salud y la vida de los pacientes, aún es necesario un adecuado proceso de validación para la implementación de estas tecnologías [7].
3. **Modelo para la detección de melanoma mediante el escaneo de imágenes con Deep Learning usando el algoritmo K-NN(Abril 2018):** El melanoma es considerado como el tumor más letal de piel por su gran capacidad de producir metástasis, a nivel mundial se estima que cada año ha sido el causante de cerca de 57,000 fallecimientos; en México, el aumento de esta neoplasia ha tenido una evolución de casi 500%, y su diagnóstico resulta de cierta forma difícil, debido a que al inicio de la neoplasia suele confundirse con verrugas. El presente trabajo de investigación proporciona una aproximación a la detección de cáncer en la piel a través de imágenes termoscópicas realizadas con una cámara web normal, utilizando el framework de Machine Learning TensorFlow como base para la creación de los bancos de datos de imágenes correspondiente, aplicando el algoritmo K-NN y el lenguaje de programación Python para su implementación y uso [8].
4. **Clasificación de cáncer de piel utilizando aprendizaje profundo:** 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, máquinas de soporte vectorial, regresión logística, entre otros; además, del uso de algoritmos de aprendizaje profundo como las redes neuronales convolucionales, ayudando a

realizar un prediagnóstico 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) \citep{ISIC:12} 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 [9].

5. **Identificación automática de cáncer de piel aplicando machine learning:** En la actualidad la implementación del machine learning en la medicina, especial en la oncología, se ha convertido en una herramienta clave para la detección de malformaciones, infecciones, problemas cardiacos, cáncer, entre otros, anomalías del cuerpo humano que al aplicar aprendizaje automático, estos procesos se convierten en herramientas que fortalecen el diagnóstico médico. En este trabajo se presenta una revisión de la literatura y los resultados obtenidos de una implementación en Python para llevar a cabo la clasificación de imágenes médicas con presencia de cáncer de piel. Se utilizaron cuatro métodos de aprendizaje distintos: máquinas de vectores de soporte, vecinos más cercanos, redes neuronales y árboles de decisión. Se utilizó la base de datos HAM10000, la cual contiene imágenes médicas manualmente etiquetadas, lo que asegura que el proceso de entrenamiento sea efectivo. En la implementación de una muestra balanceada, la que permite demostrar la eficiencia de la implementación realizada en la identificación del cáncer de piel, esta implementación tiene una gran variabilidad en los resultados, alcanzando valores cercanos al 100% de precisión en algunos casos [1].

2.4.1 MARCO LEGAL

- **Ley 1581 de 2012 (Protección de Datos Personales):** Esta ley protege la información personal de los ciudadanos colombianos, incluyendo datos relacionados con la salud. Es importante que cualquier proyecto de IA que recopile datos personales cumpla con esta ley [10].
- **Decreto 1377 de 2013:** Este decreto reglamenta la Ley 1581 de 2012 y establece los requisitos específicos para el tratamiento de datos personales [11].
- **Ley 1341 de 2009 (Ley de Transformación Digital):** Esta ley promueve la transformación digital en Colombia y establece la creación de la Agencia Nacional de Datos (AND) [12].

- **CONPES 3975 - Política Nacional de Inteligencia Artificial (2021):** Esta política establece los lineamientos para el desarrollo y la implementación de la IA en Colombia [14].

2.4.2 GLOSARIO TÉCNICO

- Aprendizaje Automático (Machine Learning): Campo de la inteligencia artificial que permite a las computadoras aprender patrones y tomar decisiones sin ser explícitamente programadas [1].
- Modelo Predictivo: Algoritmo entrenado para predecir resultados futuros basados en datos de entrada [1].
- Características (Features): Atributos o propiedades específicas de los datos de entrada utilizados por el modelo para realizar predicciones. En el caso del cáncer de piel, características podrían incluir color, tamaño, forma, entre otros aspectos de las lesiones cutáneas [6].
- Conjunto de Entrenamiento (Training Set): Conjunto de datos utilizado para entrenar el modelo de aprendizaje automático. Contiene ejemplos etiquetados de lesiones cutáneas, donde se indica si son benignas o malignas [6].
- Conjunto de Validación (Validation Set): Conjunto de datos separado del conjunto de entrenamiento, utilizado para ajustar los parámetros del modelo y evitar el sobreajuste [6].
- Conjunto de Pruebas (Test Set): Conjunto de datos independiente que se utiliza para evaluar el rendimiento del modelo después de su entrenamiento. Contiene ejemplos que el modelo no ha visto durante el proceso de entrenamiento [6].
- Sobreajuste (Overfitting): Situación en la que el modelo se ajusta demasiado a los datos de entrenamiento y tiene dificultades para generalizar a nuevos datos, lo que puede afectar negativamente su rendimiento en conjuntos de prueba [6].
- Red Neuronal Convolutiva (CNN): Tipo de red neuronal especialmente eficiente para procesar datos de rejilla, como imágenes. Es comúnmente utilizado en tareas de visión por computadora, como la clasificación de imágenes de cáncer de piel [7].
- Preprocesamiento de Datos: Etapa en la que los datos de entrada se transforman para facilitar el entrenamiento del modelo. Esto puede incluir normalización, aumento de datos y otros pasos para mejorar la calidad de los datos [7].
- Etiquetado de Datos: Proceso de asignar etiquetas a los datos de entrenamiento indicando si una lesión cutánea es benigna o maligna [7].
- Validación Cruzada (Cross-Validation): Técnica que consiste en dividir el conjunto de entrenamiento en varios subconjuntos para evaluar el modelo en múltiples combinaciones de datos de entrenamiento y validación, ayudando a obtener métricas más robustas [7].

2.5 MARCO METODOLÓGICO

La mayoría de los investigadores mencionados previamente se centran en la creación de sistemas de diagnóstico para identificar lesiones cutáneas, con un énfasis particular en el melanoma. Para ello, emplean algoritmos de inteligencia artificial que se apoyan en técnicas avanzadas de procesamiento de imágenes y aprendizaje automático para examinar y categorizar imágenes dermatoscópicas. Se han empleado diversas metodologías para llegar a estas conclusiones, entre las cuales se incluyen:

- **Procesamiento de imágenes dermatoscópicas:** Los autores emplean técnicas avanzadas de procesamiento de imágenes para mejorar la calidad de las imágenes dermatoscópicas. Esto puede incluir normalización, eliminación de ruido y realce de características [15].
- **Extracción de características:** Se aplican métodos para identificar y extraer características relevantes de las imágenes dermatoscópicas. Esto puede incluir descriptores de textura, color y bordes [16].
- **Redes neuronales profundas (Deep learning):** La mayoría de los autores hacen uso de redes neuronales profundas, especialmente redes neuronales convolucionales (CNN), que son especialmente efectivas en la clasificación de imágenes [17].
- **Entrenamiento de modelos de clasificación:** Los autores desarrollan y entrenan modelos de clasificación utilizando conjuntos de datos etiquetados. Estos modelos aprenden a reconocer patrones que permiten diferenciar entre lesiones benignas y malignas [9].
- **Integración de inteligencia artificial y expertos médicos:** Algunos autores proponen sistemas que combinan la inteligencia artificial con la evaluación de expertos médicos para mejorar la precisión del diagnóstico [18].

Después de todo lo mencionado, también es importante hacer énfasis en los algoritmos de aprendizaje supervisado que utilizaron en los diferentes proyectos, tales como:

- **K-Vecinos más cercanos (K-NN- K-Nearest Neighbors):** es un algoritmo de aprendizaje supervisado, que se encarga de analizar las características extraídas de imágenes de lesiones de piel y las clasifica como benignas o malignas, por lo que las compara con otras imágenes, si la mayoría de los vecinos pertenecen a una categoría específica (por ejemplo, maligna), entonces el sistema clasifica la lesión como

perteneciente a dicha categoría [19].

- **Arboles de decisión:** son herramientas que ayudan a clasificar datos mediante una secuencia de preguntas basadas en características específicas, se construye un árbol que toma como entrada características extraídas de imágenes de lesiones cutáneas y utiliza una serie de decisiones basadas en dichas características para determinar la clasificación de la lesión (benigna o maligna) [9].
- **Bosque aleatorios:** combinan múltiples árboles de decisión para mejorar la precisión de las predicciones, cada árbol emite su propia opinión y el resultado final se determina mediante un voto ponderado [9].
- **Regresión logística:** implica utilizar un modelo matemático que evalúa múltiples características de las lesiones cutáneas para predecir la probabilidad de que sean malignas. Se recopilan datos de lesiones previamente diagnosticadas, incluyendo factores como tamaño, forma, color y otros atributos visibles. Estos datos se utilizan para entrenar el modelo, que aprenderá a asignar pesos a cada característica en función de su relevancia para el diagnóstico. Una vez entrenado, el modelo puede evaluar nuevas lesiones y proporcionar una estimación de la probabilidad de malignidad [9].
- **Máquina de vectores de soporte:** en este caso se recopilan datos importantes de las imágenes, donde cada una cuenta con una etiqueta (maligna o benigna) y se extrae que así busque una línea o superficie que pueda separar estos dos grupos de la mejor manera posible, manteniendo la mayor distancia posible entre ellos. Esto para clasificar de manera más precisa [1].

2.5.1 POBLACIÓN Y MUESTRA

Población: Este conjunto de datos consta de 2357 imágenes de enfermedades oncológicas malignas y benignas, y fue recopilado a partir de The International Skin Imaging Collaboration (ISIC). Todas las imágenes fueron organizadas de acuerdo con la clasificación tomada por ISIC, y todos los subconjuntos se dividieron en el mismo número de imágenes, con la excepción de los melanomas y lunares, cuyas imágenes son ligeramente predominantes [14].

Muestra: Para la selección de la muestra, se aplicó la técnica de clasificación según el

sistema de ISIC. Se aseguró que cada enfermedad dentro de la población tenga una

representación equitativa en la muestra, con la excepción de los melanomas y lunares, que tienen una presencia ligeramente superior en comparación con otras enfermedades [14].

2.5.2 INSTRUMENTOS Y TÉCNICAS PARA LA RECOLECCIÓN DE LA INFORMACIÓN

Con el fin de llevar a cabo el proyecto de detección de cáncer de piel, se implementarán entrevistas y técnicas específicas para recopilar información clave. Estas estrategias son esenciales para obtener una comprensión profunda de los procesos involucrados en la detección de cáncer de piel y las necesidades específicas del sistema de aprendizaje automático.

Fuentes Primarias: La información primaria se obtendrá a través de entrevistas con expertos en el campo de la detección de cáncer de piel y el personal relacionado con el proyecto. Estas entrevistas incluirán a profesionales de la salud, dermatólogos y desarrolladores del sistema. Las reuniones se llevarán a cabo para comprender en detalle los procedimientos de detección, las prácticas médicas y las expectativas del sistema. Se prestará especial atención a las opiniones y experiencias de quienes trabajan directamente con el sistema, como ingenieros de sistemas y profesionales de la salud especializados en dermatología.

Fuentes Secundarias: Para complementar la información obtenida de las entrevistas, se utilizarán fuentes secundarias, tales como artículos científicos, proyectos similares, documentos académicos y libros relacionados con la detección de cáncer de piel mediante sistemas de aprendizaje automático. Estas fuentes servirán para respaldar y enriquecer la comprensión del estado actual de la investigación, los avances tecnológicos y las mejores prácticas en el campo de la detección de cáncer de piel.

2.6 RESULTADOS ESPERADOS

1. Selección y tratamiento del banco de datos: Se seleccionará un banco de datos apropiado que contenga imágenes de lesiones cutáneas relacionadas con el cáncer de piel, seguido

por la implementación de procesos de tratamiento de imágenes, esto incluirá la limpieza de datos y la normalización de las imágenes para prepararlas para su posterior análisis y diagnóstico.

2. Evaluación de arquitecturas de redes neuronales: Después de la exhaustiva evaluación de las diversas arquitecturas de redes neuronales de aprendizaje supervisado utilizando el banco de datos establecido, se logra la identificación de la arquitectura más adecuada para asegurar una detección precisa de cáncer de piel, basándose en métricas de rendimiento como precisión y sensibilidad.
3. Entrenamiento de la red neuronal seleccionada: La red neuronal seleccionada ha sido entrenada exitosamente utilizando el banco de datos preparado, además se han ajustado los parámetros de manera óptima para maximizar la precisión y eficiencia del modelo, siendo este logro fundamental para el rendimiento del sistema de detección.
4. Desarrollo de módulos específicos: Todos los módulos esenciales, desde la carga de datos hasta la generación de resultados, han sido desarrollados con éxito. Estos módulos aseguran un flujo eficiente de datos y el correcto funcionamiento del sistema, aspectos cruciales para la implementación exitosa del proyecto.
5. Implementación del sistema web: El sistema web para la detección de cáncer de piel ha sido implementado completamente y está disponible para su uso por parte del personal y cualquier usuario interesado. La interfaz de usuario es intuitiva, permitiendo cargar imágenes, ejecutar la detección y visualizar los resultados de manera clara y comprensible. Este logro mejora significativamente la accesibilidad y gestión de la información, facilitando la labor del personal y mejorando la experiencia de los usuarios.

**Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas**

2.7.- CRONOGRAMA DE ACTIVIDADES

Ilustración 1 Cronograma de actividades

ACTIVIDAD	ENERO				FEBRERO				MARZO				ABRIL				MAYO			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
INICIO DEL PROYECTO																				
Definir el título y objetivos del proyecto																				
DEFINICIÓN Y PLANEACIÓN																				
Realizar la revisión bibliográfica																				
Definir los objetivos del proyecto																				
Definir el alcance del proyecto																				
Construir el diseño preliminar del sistema																				
Seleccionar las herramientas y tecnologías																				
RECOPILACIÓN DE DATOS																				
Identificar fuentes de datos relevantes																				
Recopilar datos																				
Realizar el procesamiento de los datos																				
Desarrollar un conjunto de datos entrenados																				
DESARROLLO DEL MODELO DE APRENDIZAJE																				
Implementar algoritmos de aprendizaje para la detección																				
Entrenar el modelo utilizando el conjunto de datos																				
Evaluar el rendimiento del modelo																				
Realizar los ajustes necesarios																				
INTEGRACIÓN Y PRUEBAS																				
Integrar el modelo en un sistema de detección																				
Desarrollar la interfaz de usuario																				
Realizar pruebas de funcionamiento																				
OPTIMIZACIÓN, DOCUMENTACIÓN Y PRESENTACIÓN																				
Optimizar el sistema y realizar ajuste finales																				
Documentar el proyecto y los resultados obtenidos																				
Preparar la presentación final y entregar el proyecto																				

2.7 PRESUPUESTO

Tabla 1 Presupuesto de recursos físicos

RECURSOS FÍSICOS			
Recursos	Cantidad	Valor unitario(COP)	Valor total(COP)
Escritorio	2	\$204.999	\$409.800
Teclado	2	\$90.000	\$180.000
Mouse	2	\$106.700	\$213.400
Pantalla	2	\$399.000	\$798.000
Audífonos	2	\$65.000	\$130.000
Silla ergonómica	2	\$150.000	\$300.000
Base para laptop con ventilador	2	\$66.000	\$132.000
Cámara	1	\$86.900	\$86.900
TOTAL		\$1.168.599	\$2.250.100

Tabla 2: Presupuesto de recursos tecnológicos

RECURSOS TECNOLÓGICOS			
Recursos	Cantidad	Valor unitario(COP)	Valor total(COP)
Licencia de MATLAB	2	\$3.817.100	\$7.634.200
Biblioteca scikit-learn para aprendizaje automático en Python.		Gratis	\$0
Curso en línea para el tratamiento de imágenes en Udemy	2	Gratis	\$0
Curso en línea para el entrenamiento de redes neuronales con Phyton	2	\$349.000	\$698.000
Base de datos Mongo Atlas	1	\$221.886	\$1.331.316
	TOTAL	\$4.166.100	\$8.332.000

Tabla 3: Presupuesto de recursos humanos

RECURSOS HUMANOS			
Recursos	Cantidad	Valor unitario(COP)	Valor total por los 5 meses (COP)
Desarrolladores	2	\$3.500.000	\$42.000.000
	TOTAL	\$3.500.000	\$42.000.000

SECCIÓN III: DESARROLLO CIENTÍFICO – TECNOLÓGICO

3.1 – DESARROLLO DE LAS FASES DE LA METOLOGÍA PROPUESTA

3.1.1 Introducción al desarrollo metodológico

La metodología XP (Extreme Programming) ha sido seleccionada debido a su enfoque en la colaboración constante y la mejora continua, lo cual es crucial para adaptarse a los cambios y requerimientos emergentes durante el ciclo de desarrollo del proyecto de cinco meses. Este proyecto se estructuró en varias fases: definición y planificación (enero), recopilación de datos y preprocesamiento (febrero), desarrollo del modelo de aprendizaje automático (marzo), integración y pruebas del sistema (abril), y optimización, documentación y presentación (mayo).

Cada fase incluyó actividades específicas como revisión bibliográfica, definición de objetivos, preprocesamiento de datos, entrenamiento del modelo, integración del sistema, pruebas, optimización y documentación detallada. Al seguir los principios de XP, se garantizó que cada componente del sistema se desarrollara, probara y refinara de manera iterativa, asegurando la calidad y funcionalidad requerida para la detección de cáncer de piel.

3.1.1.1 Definición de roles

Para el desarrollo exitoso del sistema de aprendizaje automático para la detección de cáncer de piel, se ha formado un equipo multidisciplinario, cada miembro con roles y responsabilidades específicas. La colaboración y coordinación entre los diferentes roles es crucial para asegurar que todas las áreas del proyecto se desarrollen de manera eficiente y que los objetivos sean alcanzados conforme a los plazos establecidos. A continuación, se detallan los roles principales del equipo, junto con los encargados y una descripción de sus responsabilidades:

Roles del equipo

Tabla 4: Roles del equipo

Rol	Encargado	Descripción
Director de proyecto		Se encarga de supervisar globalmente el desarrollo del proyecto, garantizando que todos los objetivos y entregables se cumplan según el cronograma establecido. Su rol incluye coordinar eficientemente entre los miembros del equipo, asignar tareas de manera efectiva y tomar decisiones estratégicas para orientar el proyecto hacia el éxito deseado.
Desarrollador Fullstack	Stefania Scarlet Hernández Quiroz	Es responsable de la implementación tanto del frontend como del backend de una aplicación web o sistema, abarcando todas las capas del desarrollo de software. Sus tareas incluyen el diseño y la creación de la interfaz de usuario, la integración con bases de datos y servicios externos, la implementación de lógica de negocio, y la optimización del rendimiento y seguridad del sistema. del usuario en el producto final.
Desarrollador Fullstack	Alejandro Fidel Lara Miranda	Es responsable de la implementación tanto del frontend como del backend de una aplicación web o sistema, abarcando todas las capas del desarrollo de software. Sus tareas incluyen el diseño y la creación de la interfaz de usuario, la integración con bases de datos y servicios externos, la implementación de lógica de negocio, y la optimización del rendimiento y seguridad del sistema. del usuario en el producto final.

Nota: Esta tabla describe los roles y responsabilidades de los miembros del equipo.
(Elaboración propia)

3.1.1.2 Definición de responsabilidades

El desarrollo del sistema de aprendizaje automático para la detección de cáncer de piel ha sido meticulosamente organizado en diversas actividades, cada una asignada a personas específicas y enfocada en diferentes aspectos clave del proyecto. Estas actividades incluyen desde la revisión bibliográfica y la preparación de datos, hasta la implementación de algoritmos de aprendizaje automático y la integración del modelo en el sistema final. Esta estructura de trabajo asegura que todas las etapas del proyecto se desarrollen de manera coherente y eficiente, permitiendo a cada miembro del equipo concentrarse en sus áreas de especialización para alcanzar los objetivos establecidos de detección precisa y efectiva de cáncer de piel.

Listado de actividades

Tabla 5: Listado de actividades

#	Descripción de la actividad	Persona asignada
Vista principal		
1	Diseñar la interfaz de usuario para la vista skincancer, incluyendo una disposición clara y accesible de los 9 tipos de cáncer.	Alejandro Lara
2	Crear iconos y representaciones visuales para cada uno de los 9 tipos de cáncer: actinic keratosis, basal cell carcinoma, dermatofibroma, melanoma, nevus, pigmented benign keratosis, seborrheic keratosis, squamous cell carcinoma, vascular lesion.	Alejandro Lara
3	Implementar la lógica de navegación que permite al usuario seleccionar y ver detalles específicos de cada tipo de cáncer.	Alejandro Lara
Vista de entrenamiento		
4	Diseñar la interfaz de usuario para la vista de evaluación, utilizando tarjetas (cards) para representar cada uno de los 4 algoritmos: CNN, SVM, ResNet50, VGG16.	Alejandro Lara
5	Crear iconos y representaciones visuales para cada algoritmo en las tarjetas, asegurando una apariencia coherente y profesional.	Alejandro Lara
6	Implementar la lógica de selección que permite al usuario elegir uno de los algoritmos a través de las tarjetas.	Alejandro Lara
Vista de evaluación		
7	Diseñar la interfaz del módulo de evaluación, incluyendo un combo box para la selección de uno de los 4 algoritmos (CNN, SVM, ResNet50, VGG16).	Stefania Hernández

**Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas**

8	Implementar la lógica del combo box que permita al usuario seleccionar un algoritmo y actualizar la interfaz según la selección.	Stefania Hernández
9	Crear una sección para la carga de imágenes, donde los usuarios puedan subir una imagen para la predicción del cáncer de piel.	Stefania Hernández
10	Desarrollar un botón de acción que, al ser presionado, inicie el proceso de evaluación con el algoritmo seleccionado y la imagen cargada.	Stefania Hernández
Vista de resultados		
11	Diseñar la interfaz de la vista de resultados en HTML, asegurando una disposición clara y profesional.	Stefania Hernández
12	Implementar la lógica para mostrar la imagen cargada por el usuario en la vista de resultados.	Stefania Hernández
13	Desarrollar un área para mostrar el nombre del tipo de cáncer predicho basado en el análisis del algoritmo.	Stefania Hernández
14	Incluir una sección para mostrar el porcentaje de confianza del modelo en la predicción realizada.	Stefania Hernández
Modelos de aprendizaje supervisado		
15	Seleccionar y preparar los datos de entrenamiento específicos para cada algoritmo (CNN, SVM, ResNet50, VGG16).	Stefania Hernández
16	Implementar el algoritmo CNN (Convolutional Neural Network) para la detección de cáncer de piel, configurando las capas y los parámetros necesarios.	Stefania Hernández
17	Implementar el algoritmo SVM (Support Vector Machine) para la detección de cáncer de piel, ajustando el kernel y otros hiperparámetros relevantes.	Stefania Hernández
18	Implementar el algoritmo ResNet50 (Residual Networks) para la detección de cáncer de piel, importando el modelo preentrenado y ajustando las capas finales para la tarea específica.	Stefania Hernández
19	Implementar el algoritmo VGG16 para la detección de cáncer de piel, utilizando el modelo preentrenado y realizando el fine-tuning para mejorar el rendimiento en el conjunto de datos específico.	Stefania Hernández
20	Entrenar cada modelo utilizando el conjunto de datos preparado, asegurando que los modelos converjan correctamente y obteniendo las métricas de rendimiento iniciales.	Alejandro Lara
21	Evaluar el rendimiento de cada modelo en un conjunto de datos de validación, ajustando los hiperparámetros según sea necesario para mejorar la precisión y reducir el sobreajuste.	Alejandro Lara
22	Comparar los resultados de los diferentes modelos, seleccionando los mejores enfoques y documentando las razones de la elección de cada modelo.	Alejandro Lara

23	Integrar los modelos entrenados en el sistema completo de detección, asegurando que puedan ser utilizados en tiempo real para la evaluación de imágenes nuevas.	Alejandro Lara
Diseño de arquitectura		
24	Definir la arquitectura del backend, considerando aspectos como la escalabilidad, la modularidad y la seguridad.	Stefania Hernández
25	Seleccionar el modelo de arquitectura más adecuado, como arquitectura de microservicios o arquitectura monolítica, según los requisitos del proyecto	Stefania Hernández
Desarrollo de APIs		
26	Diseñar y desarrollar una API RESTful para permitir la comunicación entre el frontend y el backend de la aplicación.	Alejandro Lara
27	Definir los endpoints de la API para las diferentes funcionalidades de la aplicación	Alejandro Lara
Gestión de datos		
28	Descargar conjuntos de datos relevantes desde Kaggle utilizando la API de Kaggle o descarga directa desde la plataforma.	Stefania Hernández
29	Preparar los archivos descargados para su integración en la base de datos local del proyecto.	Stefania Hernández
30	Diseñar el esquema de la base de datos local para almacenar los datos descargados, asegurando la coherencia y estructura adecuada.	Stefania Hernández
31	Crear scripts o utilizar herramientas para cargar los datos desde los archivos descargados en la base de datos local.	Stefania Hernández

3.1.1.3 Implementación de metodología Kanban

Para gestionar las diferentes actividades del equipo en el proyecto "Sistema de aprendizaje automático para la detección de cáncer de piel", se utilizó Jira como la herramienta principal de gestión de proyectos. Jira es una plataforma en línea para el seguimiento de proyectos que permite organizar, asignar y monitorear las tareas de desarrollo de manera efectiva. En este proyecto, se adoptó la metodología Kanban, que se centra en la visualización del flujo de trabajo y la mejora continua del proceso. A continuación, se explica detalladamente cómo se implementan estas tareas en Jira, siguiendo la metodología Kanban.

3.1.1.3.1 Creación del proyecto en Jira

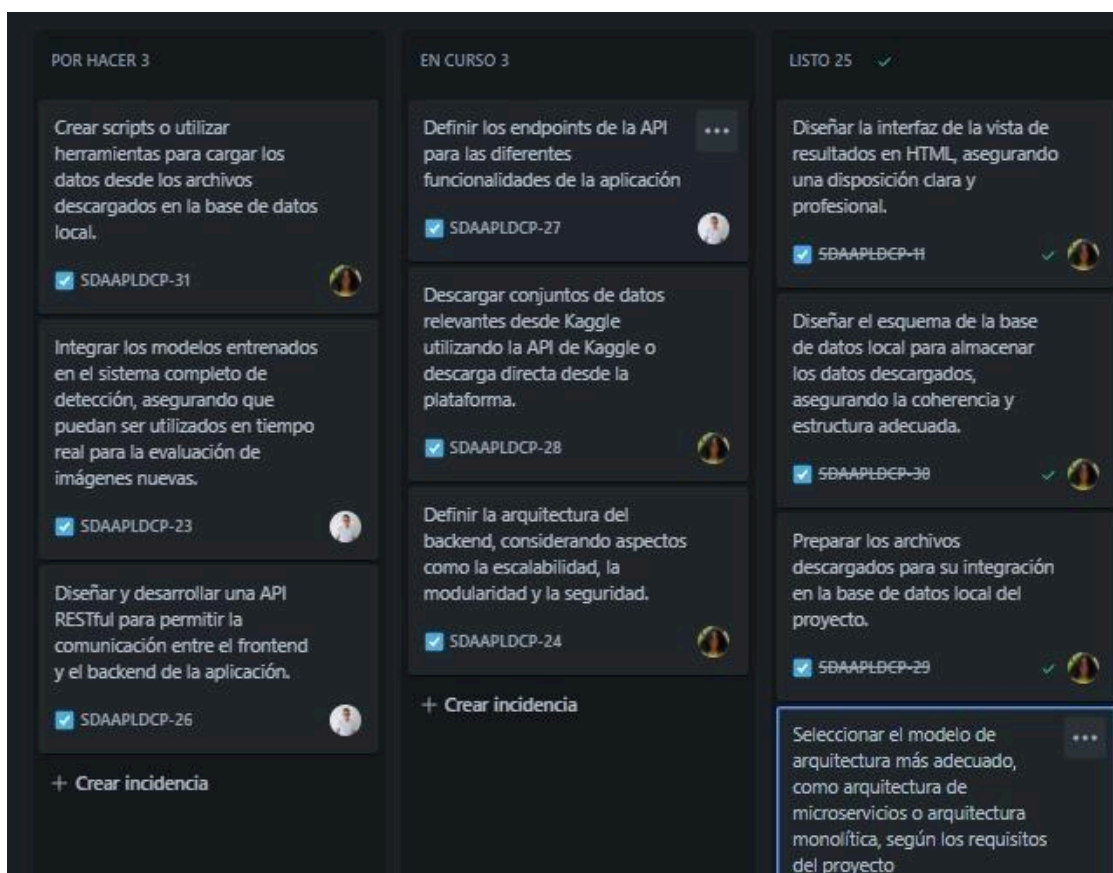
Se inició un proyecto en Jira dedicado al desarrollo del "Sistema de aprendizaje automático para la detección de cáncer de piel". El proyecto fue configurado con un tablero Kanban, perfecto para la visualización del flujo de trabajo y la gestión eficiente de las tareas.

3.1.1.3.2 Definición de columnas del tablero Kanban

Las columnas del tablero Kanban se organizaron según las diferentes fases del flujo de trabajo del proyecto. Estas columnas incluyeron:

- Pendientes: Tareas que necesitan ser priorizadas y planificadas.
- Para hacer: Tareas que están listas para ser iniciadas.
- En progreso: Tareas actualmente en desarrollo.
- Completado: Tareas que han sido finalizadas y aprobadas.

Ilustración 2: Ejemplo de tablero Kanban



Nota: Detalle de flujo de trabajo en Jira (Elaboración propia)

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

3.1.1.3.3 Creación de historias de usuario y tareas

Las actividades del proyecto se dividieron en historias de usuario y tareas específicas. Cada historia de usuario describía una funcionalidad o característica clave de la aplicación, mientras que las tareas eran los componentes específicos necesarios para completar esas historias de usuario. Por ejemplo, algunas historias de usuario incluyeron "Crear la interfaz de usuario para la vista de skincancer, asegurando una presentación clara y accesible de los 9 tipos de cáncer" y "Desarrollar la interfaz de usuario para la carga y visualización de imágenes."

3.1.1.3.4 Asignación de tareas a miembros del equipo

Cada tarea fue asignada a un miembro del equipo según sus responsabilidades y áreas de especialización, garantizando claridad en los roles y entregables. En este caso, Stefania Hernandez se encargó tanto del desarrollo frontend como backend, mientras que Alejandro Lara actuó como desarrollador fullstack, abarcando ambas áreas del proyecto.

3.1.1.3.5 Visualización del flujo de trabajo

El uso del tablero Kanban en Jira facilitó que todo el equipo pudiera visualizar de manera actualizada el estado de cada tarea. Las tarjetas de tareas se desplazaban entre las columnas conforme avanzaban en el flujo de trabajo, ofreciendo una visión clara del progreso del proyecto. En cuanto a la priorización y planificación, las tareas dentro de la columna "Backlog" fueron ordenadas según su importancia y urgencia. Durante las sesiones de planificación, las tareas de alta prioridad se movieron a la columna "Por Hacer" para su ejecución.

3.1.1.3.6 Ejecución de tareas

Los miembros del equipo tomaban tareas de la columna "Por Hacer" y las movían a "En Progreso" cuando empezaban a trabajar en ellas. Esto permitía mantener un registro claro de quién estaba encargado de cada tarea y en qué fase del desarrollo se encontraba.

3.1.1.3.7 Revisión y aprobación

Al completar una tarea, esta se trasladaba a la columna "En Revisión". En esta fase, otros miembros del equipo revisaban el trabajo para asegurar su calidad y que cumpliera con los

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

requisitos establecidos. Las tareas que pasaban la revisión se movían a "Completadas", cerrando así el ciclo de desarrollo.

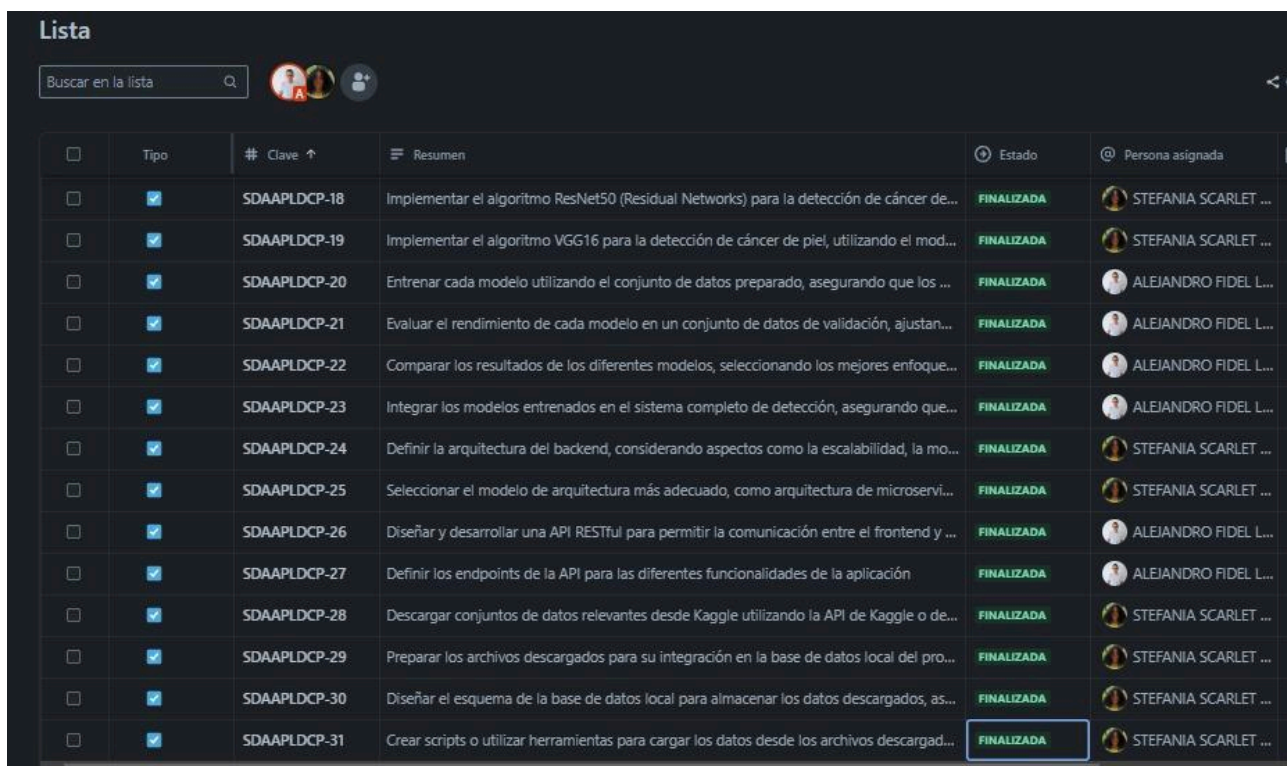
3.1.1.3.8 Comentarios y colaboración

Los miembros del equipo podían añadir comentarios en las tarjetas de tareas para discutir detalles, formular preguntas o proporcionar actualizaciones. Esta funcionalidad de colaboración fue crucial para mantener una comunicación clara y continua.

3.1.1.3.9 Seguimiento del progreso

Jira ofrecía informes y gráficos (como el diagrama de flujo acumulado y el diagrama de control) que ayudaban a monitorear el progreso del proyecto y a identificar posibles cuellos de botella en el flujo de trabajo.

Ilustración 3: Cronograma de actividades Jira



<input type="checkbox"/>	Tipo	# Clave ↑	Resumen	Estado	Persona asignada
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-18	Implementar el algoritmo ResNet50 (Residual Networks) para la detección de cáncer de...	FINALIZADA	STEFANIA SCARLET ...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-19	Implementar el algoritmo VGG16 para la detección de cáncer de piel, utilizando el mod...	FINALIZADA	STEFANIA SCARLET ...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-20	Entrenar cada modelo utilizando el conjunto de datos preparado, asegurando que los ...	FINALIZADA	ALEJANDRO FIDEL L...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-21	Evaluar el rendimiento de cada modelo en un conjunto de datos de validación, ajustan...	FINALIZADA	ALEJANDRO FIDEL L...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-22	Comparar los resultados de los diferentes modelos, seleccionando los mejores enfoque...	FINALIZADA	ALEJANDRO FIDEL L...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-23	Integrar los modelos entrenados en el sistema completo de detección, asegurando que...	FINALIZADA	ALEJANDRO FIDEL L...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-24	Definir la arquitectura del backend, considerando aspectos como la escalabilidad, la mo...	FINALIZADA	STEFANIA SCARLET ...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-25	Seleccionar el modelo de arquitectura más adecuado, como arquitectura de microservi...	FINALIZADA	STEFANIA SCARLET ...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-26	Diseñar y desarrollar una API RESTful para permitir la comunicación entre el frontend y ...	FINALIZADA	ALEJANDRO FIDEL L...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-27	Definir los endpoints de la API para las diferentes funcionalidades de la aplicación	FINALIZADA	ALEJANDRO FIDEL L...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-28	Descargar conjuntos de datos relevantes desde Kaggle utilizando la API de Kaggle o de...	FINALIZADA	STEFANIA SCARLET ...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-29	Preparar los archivos descargados para su integración en la base de datos local del pro...	FINALIZADA	STEFANIA SCARLET ...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-30	Diseñar el esquema de la base de datos local para almacenar los datos descargados, as...	FINALIZADA	STEFANIA SCARLET ...
<input type="checkbox"/>	<input checked="" type="checkbox"/>	SDAAPLDCP-31	Crear scripts o utilizar herramientas para cargar los datos desde los archivos descargad...	FINALIZADA	STEFANIA SCARLET ...

Nota: Esta figura muestra las actividades de los diversos módulos que se manejan en el aplicativo. (Elaboración propia)

La implementación de Jira con la metodología Kanban permitió una gestión eficiente y visual del proyecto, asegurando que todas las tareas se completaran de manera ordenada y dentro de los plazos establecidos. La capacidad de visualizar el flujo de trabajo, priorizar tareas, colaborar efectivamente y seguir el progreso en tiempo real contribuyó significativamente al éxito del desarrollo de la aplicación. Esta estructura organizada y detallada facilitó la coordinación entre los miembros del equipo y garantizó que cada etapa del proyecto se ejecutara con la máxima eficiencia y calidad.

3.1.2 Etapa de investigación

Para investigar exhaustivamente el estado actual y la efectividad del "Sistema de aprendizaje automático para la detección de cáncer de piel", se llevó a cabo un estudio multidimensional que incluyó diversas fuentes de información y métodos de recolección de datos. A continuación, se detalla el proceso seguido y los hallazgos clave obtenidos:

3.1.2.1 Fuentes de investigación

1. **Bases de datos públicas de imágenes médicas:** Se utilizaron bases de datos como la ISIC (International Skin Imaging Collaboration), que proporcionaron un conjunto extenso y variado de imágenes de cáncer de piel. Esta fuente fue crucial para obtener datos oficiales y recientes.
2. **Entrevistas con ingenieros de sistemas:** Se realizaron entrevistas estructuradas con ingenieros de sistemas que han trabajado con algoritmos de aprendizaje automático en el ámbito de la detección de cáncer de piel. Las preguntas abarcaron temas como la implementación de modelos, los desafíos técnicos encontrados y las mejores prácticas para optimizar el rendimiento de los algoritmos.
3. **Revisión de literatura científica:** Se llevó a cabo una revisión exhaustiva de artículos científicos y estudios recientes sobre el uso de aprendizaje automático en la detección de cáncer de piel. Estos documentos proporcionaron información valiosa sobre las últimas técnicas, resultados de investigaciones previas y las mejores prácticas en el campo.

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

3.1.2.2 Desarrollo de la investigación preliminar

3.1.2.2.1 Bases de datos públicas de imágenes médicas

Descripción: Se utilizaron bases de datos como la ISIC (International Skin Imaging Collaboration), que proporcionan un extenso conjunto de imágenes de cáncer de piel, incluyendo diferentes tipos y estadios de la enfermedad.

Proceso de recolección: Se accedió a estas bases de datos mediante plataformas en línea, descargando imágenes relevantes para entrenar y evaluar los modelos de aprendizaje automático.

Importancia: Estas bases de datos son cruciales ya que ofrecen imágenes etiquetadas por expertos, lo que es esencial para la precisión y confiabilidad del entrenamiento de los modelos. Además, la diversidad de las imágenes permite una mejor generalización del modelo a diferentes casos y condiciones.

3.1.2.2.2 Entrevistas con ingenieros de sistemas

Descripción: Se realizaron entrevistas estructuradas con ingenieros de sistemas que tienen experiencia en la implementación y optimización de algoritmos de aprendizaje automático para la detección de cáncer de piel.

Proceso de recolección: Las entrevistas se llevaron a cabo de manera presencial y virtual, utilizando un cuestionario diseñado para abordar aspectos técnicos y prácticos de la implementación de estos algoritmos.

Temas abordados: Las entrevistas cubrieron temas como la selección de modelos, los desafíos técnicos encontrados durante el desarrollo, y las estrategias utilizadas para mejorar la precisión y eficiencia de los algoritmos.

Importancia: Estas entrevistas proporcionaron información valiosa sobre las mejores prácticas y desafíos reales enfrentados durante el desarrollo de sistemas de detección de cáncer de piel, ofreciendo insights prácticos y aplicables al proyecto.

3.1.2.2.3 Revisión de literatura científica

Descripción: Se realizó una revisión exhaustiva de artículos científicos y estudios recientes relacionados con el uso de aprendizaje automático en la detección de cáncer de piel.

Proceso de recolección: Los artículos se obtuvieron a través de bases de datos académicas como PubMed, IEEE Xplore, y Google Scholar. Se seleccionaron estudios relevantes basados en criterios de calidad y relevancia.

**Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas**

Temas abordados: La revisión incluyó el análisis de diferentes enfoques de modelado, técnicas de preprocesamiento de imágenes, métodos de validación de modelos, y resultados de estudios previos.

Importancia: La revisión de literatura proporcionó una base teórica sólida, permitiendo identificar las tendencias actuales, los avances más recientes, y las lagunas en la investigación, lo que ayudó a orientar el desarrollo del proyecto.

3.1.2.3 Recolección de información

3.1.2.3.1 Diseño del formulario de preguntas

Para obtener datos detallados y precisos sobre el uso y la percepción del sistema de detección de cáncer de piel, se diseñó un formulario de preguntas dirigido a ingenieros de sistemas. Este formulario está estructurado para capturar tanto evaluaciones cuantitativas como cualitativas de los usuarios, proporcionando una base sólida para análisis y mejoras futuras.

Tabla 6: Formulario de recolección de información

N°	Pregunta	Propósito
1	¿Con qué frecuencia utiliza el sistema de detección de cáncer de piel?	Determinar el nivel de uso del sistema entre los ingenieros de sistemas. Esto ayudará a evaluar la relevancia y la utilidad del sistema en la práctica diaria de los usuarios.
2	¿Qué tan familiarizado está con los algoritmos de entrenamiento CNN utilizados en el desarrollo del sistema?	Medir el grado de conocimiento y experiencia de los ingenieros con los algoritmos CNN, lo que puede influir en su percepción y uso del sistema.
3	¿Qué tan familiarizado está con los algoritmos de entrenamiento SVM utilizados en el desarrollo del sistema?	Evaluar el conocimiento y la experiencia de los ingenieros con los algoritmos SVM, y entender cómo esto puede afectar su confianza en el sistema.
4	¿Qué tan familiarizado está con los algoritmos de entrenamiento VGG16 utilizados en el desarrollo del sistema?	Determinar el nivel de familiaridad de los ingenieros con VGG16, un algoritmo específico, y su impacto en la aceptación del sistema.
5	¿Qué tan familiarizado está con los algoritmos de entrenamiento ResNet50 utilizados en el desarrollo del sistema?	Conocer el grado de conocimiento de los ingenieros sobre ResNet50 y cómo esto puede influir en su percepción de la efectividad del sistema.

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

6	¿Cómo se enteró del sistema de detección de cáncer de piel?	Identificar las principales fuentes de información que llevan a los ingenieros a conocer y utilizar el sistema, lo que puede orientar estrategias de difusión y marketing.
7	¿Qué tan útil encuentra el sistema para su trabajo diario?	Evaluar la percepción de utilidad del sistema en las actividades diarias de los ingenieros, lo que puede indicar su valor práctico y relevancia
8	¿Qué tan preciso considera que es el sistema en la detección de cáncer de piel?	Obtener una evaluación de la precisión del sistema desde la perspectiva de los usuarios, lo que es crucial para validar su efectividad y confiabilidad.
9	¿Qué tan fácil de usar encuentra el sistema?	Medir la facilidad de uso del sistema, un factor importante para su adopción y satisfacción del usuario.
10	¿Ha utilizado alguno de los siguientes algoritmos en sus propios proyectos?	Identificar los algoritmos más comunes utilizados por los ingenieros, lo que puede correlacionarse con su nivel de experiencia y confianza en el sistema.
11	En su experiencia, ¿Cuál de estos algoritmos considera más efectivo para la detección de cáncer de piel?	Recopilar opiniones sobre la efectividad comparativa de diferentes algoritmos, lo que puede guiar mejoras futuras en el sistema.
12	¿Recomendaría este sistema a otros profesionales?	Evaluar el nivel de satisfacción general y la disposición de los ingenieros a recomendar el sistema, lo que es un indicador clave de su éxito y aceptación en la comunidad profesional.

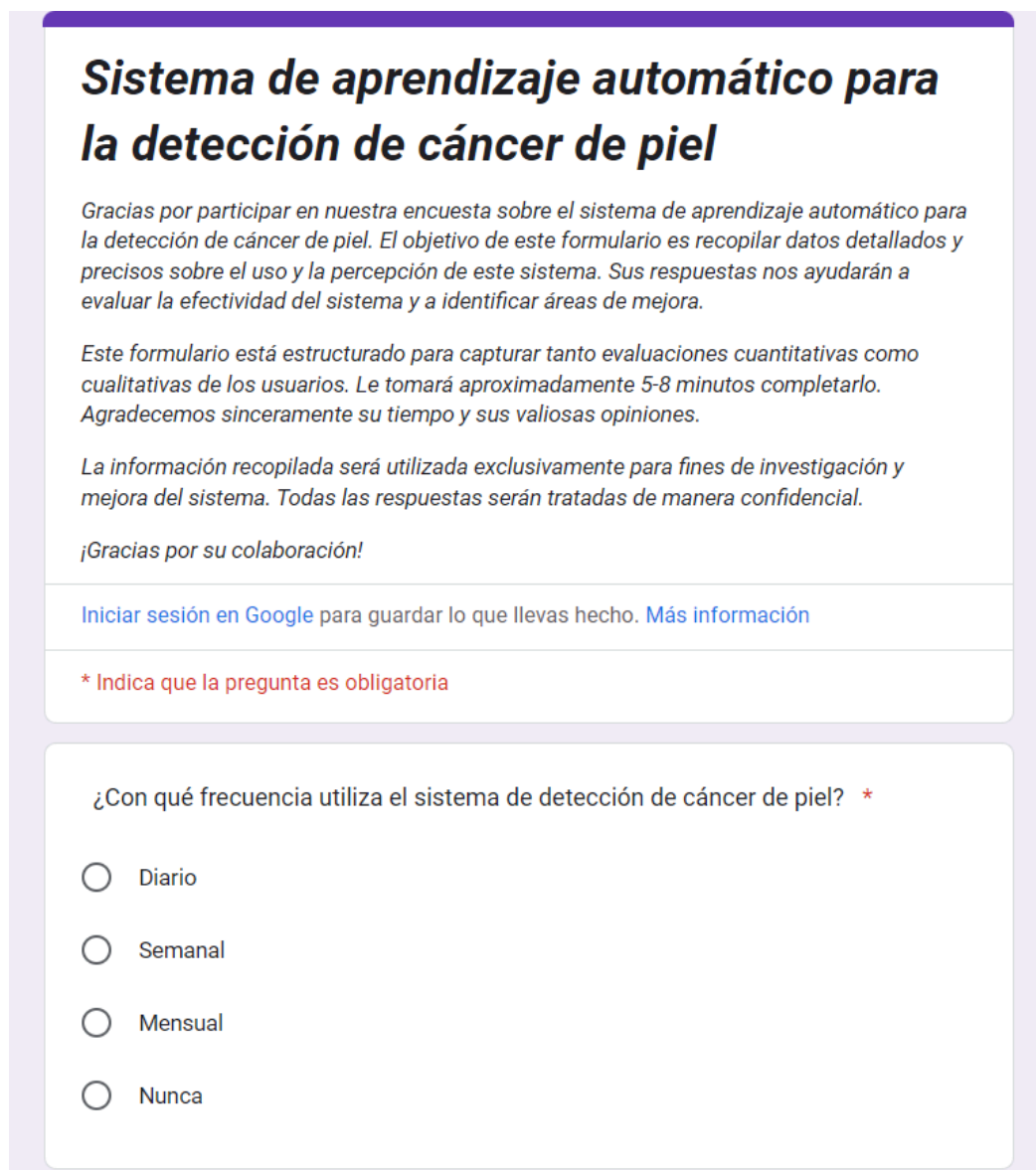
Nota: Esta tabla describe las preguntas para la recolección de información. (Elaboración propia)

3.1.2.3.2 Implementación del formulario

1. Plataforma de implementación: Se utilizó Google Forms para la implementación del formulario debido a su accesibilidad y facilidad de uso. Esta herramienta permite una distribución digital sencilla a través de enlaces directos, redes sociales o integraciones en sitios web, facilitando así el acceso a una amplia gama de usuarios.

2. Método de distribución: El formulario se compartió mediante plataformas de redes sociales y correo electrónico, con el objetivo de alcanzar un número significativo de interacciones por parte de los usuarios.

Ilustración 4: Encuesta para la recolección de información



Sistema de aprendizaje automático para la detección de cáncer de piel

Gracias por participar en nuestra encuesta sobre el sistema de aprendizaje automático para la detección de cáncer de piel. El objetivo de este formulario es recopilar datos detallados y precisos sobre el uso y la percepción de este sistema. Sus respuestas nos ayudarán a evaluar la efectividad del sistema y a identificar áreas de mejora.

Este formulario está estructurado para capturar tanto evaluaciones cuantitativas como cualitativas de los usuarios. Le tomará aproximadamente 5-8 minutos completarlo. Agradecemos sinceramente su tiempo y sus valiosas opiniones.

La información recopilada será utilizada exclusivamente para fines de investigación y mejora del sistema. Todas las respuestas serán tratadas de manera confidencial.

¡Gracias por su colaboración!

[Iniciar sesión en Google](#) para guardar lo que llevas hecho. [Más información](#)

* Indica que la pregunta es obligatoria

¿Con qué frecuencia utiliza el sistema de detección de cáncer de piel? *

- Diario
- Semanal
- Mensual
- Nunca

Nota: Esta figura evidencia el formulario empleado para la recolección de información.
(Elaboración propia)

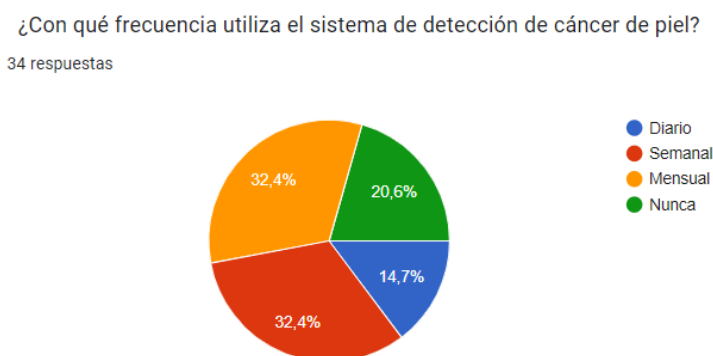
3.1.2.4 Resultados obtenidos

3.1.2.4.1 Herramientas de análisis

Para el análisis de los datos recolectados mediante el formulario, se utilizaron Microsoft Excel y las herramientas integradas de Google Forms. Excel proporciona herramientas avanzadas para realizar análisis descriptivo, incluyendo la generación de tablas de frecuencia, gráficos de barras y de pastel, así como el cálculo de medidas de tendencia central. Estas funcionalidades son cruciales para interpretar los datos de manera efectiva y obtener conclusiones significativas.

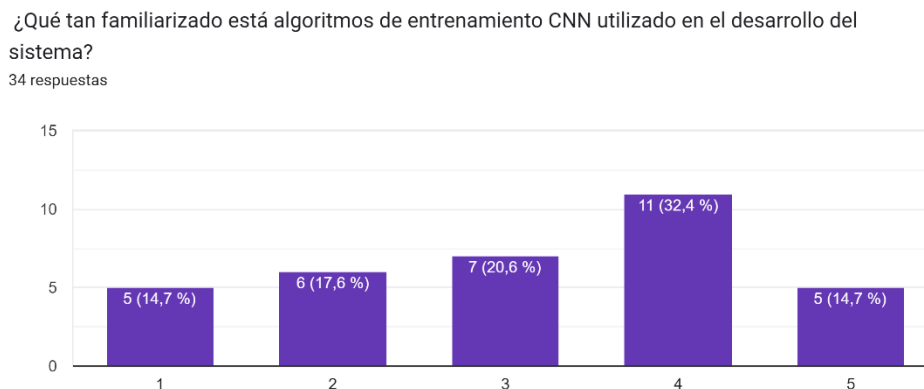
Presentación de resultados y gráficos

Ilustración 5: Gráfica de resultados del formulario de recolección de datos 1



Nota: Esta figura representa en porcentaje los resultados obtenidos ante la primera pregunta del formulario de recolección de datos.

Ilustración 6: Gráfica de resultados del formulario de recolección de datos 2



Nota: Esta figura representa en porcentaje los resultados obtenidos ante la segunda

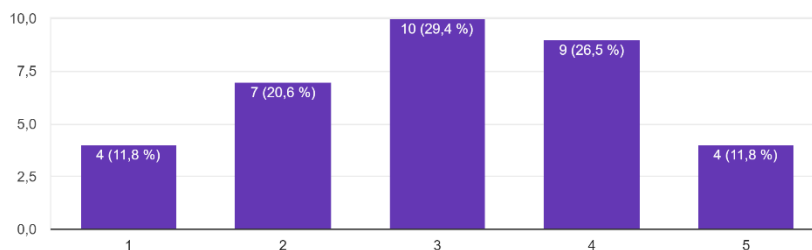
Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

pregunta del formulario de recolección de datos.

Ilustración 7: Gráfica de resultados del formulario de recolección de datos 3

¿Qué tan familiarizado está algoritmos de entrenamiento SVM utilizado en el desarrollo del sistema?

34 respuestas

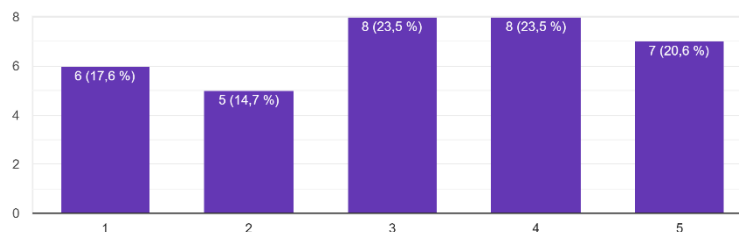


Nota: Esta figura representa en porcentaje los resultados obtenidos ante la tercera pregunta del formulario de recolección de datos.

Ilustración 8: Gráfica de resultados del formulario de recolección de datos 4

¿Qué tan familiarizado está algoritmos de entrenamiento VGG16 utilizado en el desarrollo del sistema?

34 respuestas

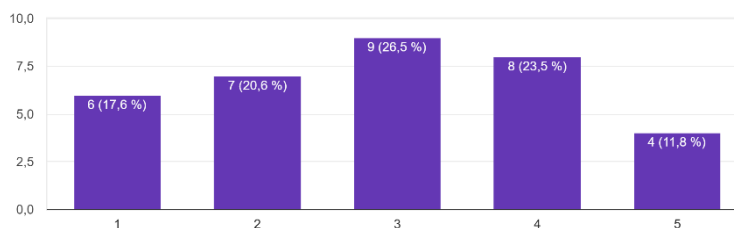


Nota: Esta figura representa en porcentaje los resultados obtenidos ante la cuarta pregunta del formulario de recolección de datos.

Ilustración 9: Gráfica de resultados del formulario de recolección de datos 5

¿Qué tan familiarizado está algoritmos de entrenamiento ResNet50 utilizado en el desarrollo del sistema?

34 respuestas



Nota: Esta figura representa en porcentaje los resultados obtenidos ante la quinta pregunta

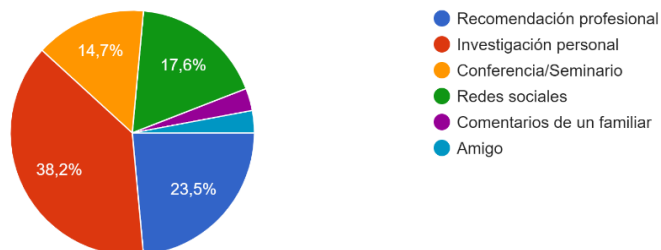
Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

del formulario de recolección de datos.

Ilustración 10: Gráfica de resultados del formulario de recolección de datos 6

¿Cómo se enteró del sistema de detección de cáncer de piel?

34 respuestas

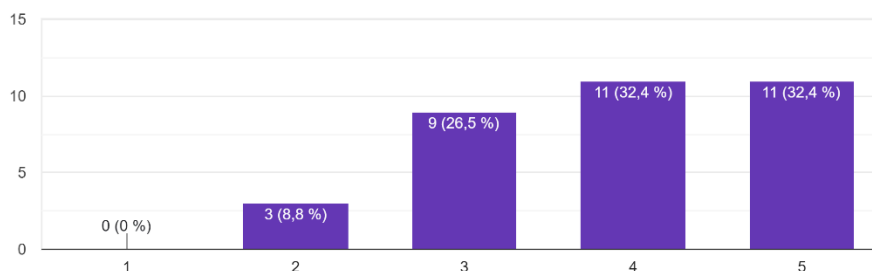


Nota: Esta figura representa en porcentaje los resultados obtenidos ante la sexta pregunta del formulario de recolección de datos.

Ilustración 11: Gráfica de resultados del formulario de recolección de datos 7

¿Qué tan útil encuentra el sistema para su trabajo diario?

34 respuestas

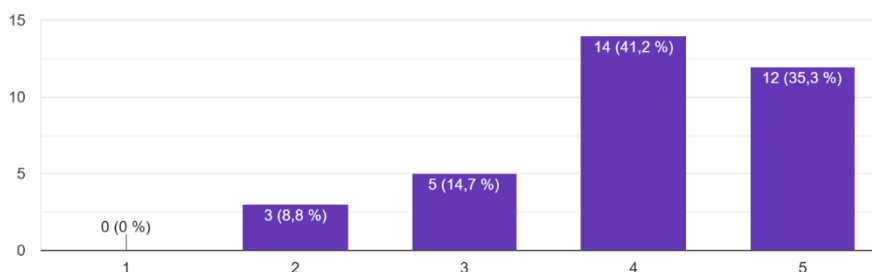


Nota: Esta figura representa en porcentaje los resultados obtenidos ante la séptima pregunta del formulario de recolección de datos.

Ilustración 12: Gráfica de resultados del formulario de recolección de datos 8

¿Qué tan preciso considera que es el sistema en la detección de cáncer de piel?

34 respuestas



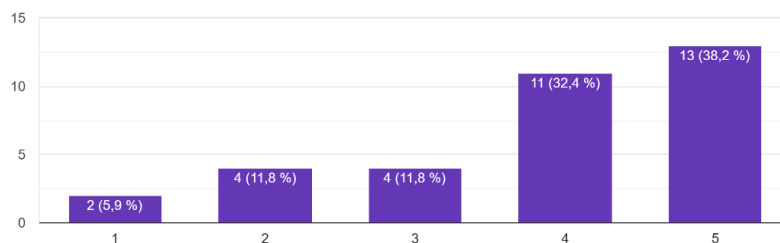
Nota: Esta figura representa en porcentaje los resultados obtenidos ante la octava pregunta

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

del formulario de recolección de datos.

Ilustración 13: Gráfica de resultados del formulario de recolección de datos 9

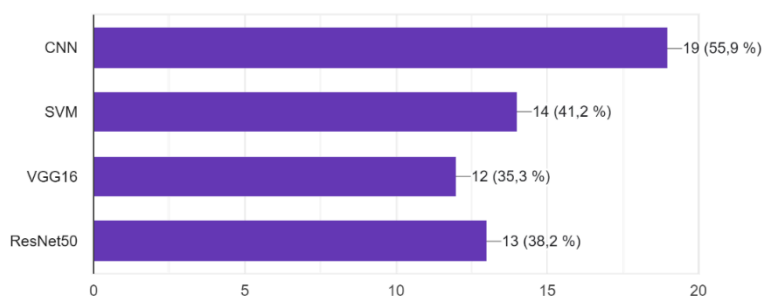
¿Qué tan fácil de usar encuentra el sistema?
34 respuestas



Nota: Esta figura representa en porcentaje los resultados obtenidos ante la novena pregunta del formulario de recolección de datos.

Ilustración 14: Gráfica de resultados del formulario de recolección de datos 10

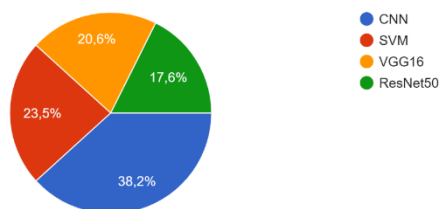
¿Ha utilizado alguno de los siguientes algoritmos en sus propios proyectos?
34 respuestas



Nota: Esta figura representa en porcentaje los resultados obtenidos ante la décima pregunta del formulario de recolección de datos.

Ilustración 15: Gráfica de resultados del formulario de recolección de datos 11

En su experiencia, ¿Cuál de estos algoritmos considera más efectivo para la detección de cáncer de piel?
34 respuestas



Nota: Esta figura representa en porcentaje los resultados obtenidos ante la once pregunta

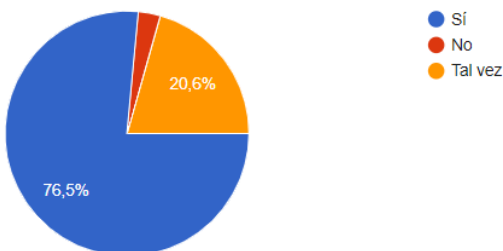
Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

del formulario de recolección de datos.

Ilustración 16: Gráfica de resultados del formulario de recolección de datos 12

¿Recomendaría este sistema a otros profesionales?

34 respuestas



Nota: Esta figura representa en porcentaje los resultados obtenidos ante la doce pregunta del formulario de recolección de datos.

3.1.2.5 Análisis de resultados

Basado en los resultados obtenidos de la encuesta realizada a 34 personas sobre el sistema de aprendizaje automático de cáncer de piel, se puede extraer varias conclusiones claves que demuestran la necesidad y aceptación potencial, a continuación se presenta un análisis detallado:

Tabla 7 Preguntas de la encuesta que se realizo

Pregunta	Aspecto	Porcentaje
¿Con qué frecuencia utiliza el sistema de detección de cáncer de piel?	Diario	14,7 %
	Semanal	32,4 %
	Mensual	32,4 %
	Nunca	20,6 %
¿Qué tan familiarizado está algoritmos de entrenamiento CNN utilizado en el desarrollo del sistema?	Desconocido	14,7 %
	Poco conocido	17,6 %
	Moderadamente conocido	20,6 %
	Conocido	32,4 %
	Muy conocido	14,7 %
¿Qué tan familiarizado está algoritmos de entrenamiento SVM utilizado en el desarrollo del sistema?	Desconocido	11,8 %
	Poco conocido	20,6 %
	Moderadamente conocido	29,4 %
	Conocido	26,5 %

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

	Muy conocido	11,8 %
--	--------------	--------

¿Qué tan familiarizado está algoritmos de entrenamiento VGG16 utilizado en el desarrollo del sistema?	Desconocido	17,6 %
	Poco conocido	14,7 %
	Moderadamente conocido	23,5 %
	Conocido	23,5 %
	Muy conocido	20,6 %
¿Qué tan familiarizado está algoritmos de entrenamiento ResNet50 utilizado en el desarrollo del sistema?	Desconocido	17,6 %
	Poco conocido	20,6 %
	Moderadamente conocido	26,5 %
	Conocido	23,5 %
	Muy conocido	11,8 %
¿Cómo se enteró del sistema de detección de cáncer de piel?	Recomendación profesional	23,5 %
	Investigación personal	38,2 %
	Conferencia/Seminario	14,7 %
	Redes sociales	17,6 %
	Comentarios de un familiar	2,9 %
	Amigo	2,9 %
¿Qué tan útil encuentra el sistema para su trabajo diario?	Nada útil	0 %
	Poco útil	8,8 %
	Moderadamente útil	26,5 %
	Útil	32,4 %
	Muy útil	32,4 %
¿Qué tan preciso considera que es el sistema en la detección de cáncer de piel?	Nada preciso	0 %
	Poco preciso	8,8 %
	Moderadamente preciso	14,7 %
	Preciso	41,2 %
	Muy preciso	35,3 %
¿Qué tan fácil de usar encuentra el sistema?	Nada fácil	5,9 %
	Poco fácil	11,8 %
	Moderadamente fácil	11,8 %
	Fácil	32,4 %
	Muy fácil	38,2 %
¿Ha utilizado alguno de los siguientes algoritmos en sus propios proyectos?	CNN	55,9 %
	SVM	41,2 %
	VGG16	35,3 %
	ResNet50	38,2 %

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

En su experiencia, ¿Cuál de estos algoritmos considera más efectivo para la detección de cáncer de piel?	CNN	38,2 %
	SVM	23,5 %
	VGG16	20,6 %
	ResNet50	17,6 %
¿Recomendaría este sistema a otros profesionales?	Sí	76,5 %
	No	20,6 %
	Tal vez	2,9 %

Nota: Esta tabla detalla las preguntas del cuestionario utilizado para la recolección de información sobre las necesidades actuales de los usuarios con respecto al sistema de aprendizaje automático para la detección de cáncer de piel (Elaboración Propia)

Los resultados de la encuesta revelan varios aspectos importantes sobre la percepción y uso del sistema de detección de cáncer de piel. En primer lugar, existe una adopción moderada del sistema, con un 32,4% de los usuarios utilizándolo semanal o mensualmente, lo que indica que es una herramienta recurrente en su práctica profesional. Sin embargo, un 20,6% nunca lo ha utilizado, lo que sugiere que aún hay margen para aumentar su adopción.

En cuanto al conocimiento sobre los algoritmos utilizados, la mayoría de los usuarios están familiarizados con CNN, SVM, VGG16 y ResNet50, pero el nivel de familiaridad varía, siendo CNN el más conocido (32,4% lo considera "conocido"). Esto resalta la importancia de ofrecer más formación y recursos para ayudar a los usuarios a comprender mejor estos algoritmos y sus aplicaciones.

Además, el sistema es percibido como útil y preciso, con un 32,4% de los encuestados considerándolo "muy útil" y un 35,3% "muy preciso". Estos resultados son alentadores, ya que reflejan la confianza de los usuarios en la herramienta. La facilidad de uso también es destacada, con un 38,2% considerándola "muy fácil" de usar. La recomendación del sistema por parte del 76,5% de los usuarios es un indicador positivo de la satisfacción general y la efectividad percibida del sistema.

En conclusión, la encuesta subraya la importancia de seguir mejorando el sistema, especialmente en áreas de formación sobre los algoritmos utilizados y en estrategias para aumentar su adopción. Además, confirma que el sistema es una herramienta valiosa y

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

efectiva para los profesionales, lo que justifica su promoción y mejora continua.

3.1.3 Preparación y planificación

En esta etapa del proyecto, se analiza la investigación realizada sobre el estado actual de los sistemas de detección de cáncer de piel mediante aprendizaje automático para comprender mejor la tecnología existente y las necesidades de los profesionales de la salud. A partir de este análisis, se definen los requerimientos específicos del sistema, tanto funcionales como no funcionales, que permitieron diseñar soluciones efectivas y pertinentes para mejorar la precisión y facilidad de uso del sistema de detección. Las historias de usuario fueron desarrolladas para cada requerimiento funcional, proporcionando un marco claro y estructurado para el desarrollo y pruebas del sistema, asegurando que todas las funcionalidades respondan de manera efectiva a las necesidades reales de los usuarios en la detección temprana de cáncer de piel.

3.1.3.1 Requerimientos funcionales

Los requerimientos funcionales especifican las funcionalidades esenciales que el sistema debe proporcionar a los usuarios. Estas funcionalidades definen las interacciones directas del usuario con el sistema, cada una diseñada para cumplir con un aspecto específico del proceso de detección de cáncer de piel, facilitando así una experiencia de usuario fluida y eficiente. Estos incluyen:

Tabla 8 Requerimientos funcionales

REF	NOMBRE	DESCRIPCIÓN	PRIORIDAD
RF01	Módulo de inicio	El sistema debe mostrar información detallada sobre 9 tipos diferentes de cáncer de piel.	Alta
RF02	Módulo de inicio	El sistema debe permitir a los usuarios navegar entre los diferentes tipos de cáncer de piel.	Alta
RF03	Módulo de entrenamiento	El sistema debe permitir al usuario seleccionar entre 4 tipos de algoritmos de entrenamiento: SVM, VGG16, CNN, y ResNet50.	Alta
RF04	Módulo de entrenamiento	El sistema debe ofrecer un botón "Entrenar" para cada uno de los 4 algoritmos, que inicie el proceso de	Alta

Universidad Popular del Cesar
Departamento de Ingeniería de
Sistemas

		entrenamiento del modelo.	
--	--	---------------------------	--

RF05	Módulo de entrenamiento	El sistema debe mostrar una imagen representativa del modelo entrenado tras finalizar el proceso de entrenamiento.	Media
RF06	Módulo de entrenamiento	El sistema debe permitir evaluar si una imagen corresponde a un melanoma u otro tipo de cáncer, basado en el modelo seleccionado por el usuario.	Alta
RF07	Módulo de Evaluación	El sistema debe permitir al usuario cargar una imagen desde su dispositivo para evaluación.	Alta
RF08	Módulo de Evaluación	El sistema debe permitir al usuario seleccionar uno de los 4 algoritmos (SVM, VGG16, CNN, ResNet50) para evaluar la imagen cargada.	Alta
RF09	Módulo de Evaluación	El sistema debe mostrar el resultado de la evaluación, indicando si la imagen corresponde a un melanoma u otro tipo de cáncer de piel, junto con un porcentaje de confianza en la predicción.	Alta

Nota: Esta tabla muestra los requerimientos funcionales considerados para el presente proyecto.

(Elaboración propia)

3.1.3.2 Requerimientos no funcionales

Los requerimientos no funcionales describen las cualidades y atributos del sistema, tales como la disponibilidad, la seguridad y la eficiencia, que son fundamentales para apoyar el funcionamiento del sistema sin especificar directamente las actividades realizadas por el usuario. Estos requerimientos son cruciales para asegurar la operatividad, la eficacia y la seguridad del sistema en su conjunto. A continuación, se detallan los requerimientos no funcionales a tener en cuenta:

Tabla 9 Requerimientos no funcionales

REF	NOMBRE	DESCRIPCIÓN
RNF01	Disponibilidad	El sistema debe mostrar información detallada sobre 9 tipos diferentes de cáncer de piel. El sistema debe estar disponible para los usuarios el 99.9% del tiempo, asegurando un acceso ininterrumpido durante los periodos de uso previstos, especialmente en entornos clínicos y de diagnóstico.
RNF02	Seguridad	El sistema debe implementar medidas de seguridad robustas para proteger la confidencialidad e integridad de los datos del usuario, incluyendo la información personal y las imágenes médicas cargadas para la evaluación.
RNF03	Eficiencia	El sistema debe procesar y entregar los resultados de las evaluaciones en un tiempo máximo de 5 segundos por imagen, garantizando una respuesta rápida y eficiente que no comprometa la calidad del diagnóstico.
RNF04	Escalabilidad	El sistema debe ser escalable, permitiendo la adición de nuevas funcionalidades o el soporte para un mayor número de usuarios y volúmenes de datos sin una disminución perceptible en el rendimiento.
RNF05	Usabilidad	El sistema debe ser fácil de usar y comprender, con una interfaz de usuario intuitiva que minimice la necesidad de formación adicional para los profesionales de la salud que lo utilicen.
RNF06	Compatibilidad	El sistema debe ser compatible con una variedad de dispositivos y sistemas operativos, incluidos Windows, macOS, y Linux, así como navegadores webs comunes como Chrome, Firefox, y Safari.
RNF07	Mantenibilidad	El sistema debe ser fácilmente mantenible, permitiendo actualizaciones y correcciones de errores sin interrumpir su operatividad ni la experiencia del usuario.
RNF08	Rendimiento	El sistema debe manejar cargas de trabajo variables sin degradar su rendimiento, asegurando una experiencia de usuario consistente incluso durante picos de uso.

Nota: Esta tabla muestra los requerimientos no funcionales considerados para el presente proyecto. (Elaboración propia)

Los requerimientos no funcionales describen las cualidades y atributos del sistema, tales como la disponibilidad, la seguridad y la eficiencia, que son fundamentales para apoyar el funcionamiento del sistema sin especificar directamente las actividades realizadas por el usuario. Estos requerimientos son cruciales para asegurar la operatividad, la eficacia y la seguridad del sistema en su conjunto. A continuación, se detallan los requerimientos no funcionales a tener en cuenta:

Tabla 10 Historia de usuario: Visualización de información sobre tipos de cáncer de piel

REF: HU01	Actor: Usuario
Nombre de la Historia	Visualización de información sobre tipos de cáncer de piel
Prioridad	Alta
Tiempo estimado	4 horas
Responsable	Alejandro Fidel Lara Miranda
Descripción	Como usuario, quiero visualizar información detallada sobre los 9 tipos de cáncer de piel en la página de inicio, para comprender mejor las diferentes formas de esta enfermedad.
Criterios de aceptación	La información es fácilmente accesible y navegable desde la página de inicio.

Nota: Esta tabla describe la historia de usuario relacionada a la funcionalidad de “Visualización de información sobre tipos de cáncer de piel”. (Elaboración propia)

Tabla 11 Historia de usuario: Selección de algoritmo para entrenamiento

REF: HU02	Actor: Usuario
Nombre de la Historia	Selección de algoritmo para entrenamiento
Prioridad	Alta
Tiempo estimado	6 horas
Responsable	Stefania Hernández Quiroz
Descripción	Como usuario, quiero poder seleccionar entre los algoritmos SVM, VGG16, CNN, y ResNet50 para entrenar el modelo, para poder elegir el algoritmo más adecuado para mi análisis.
Criterios de aceptación	El entrenamiento se completa sin errores y el sistema confirma la finalización del proceso.

Nota: Esta tabla describe la historia de usuario relacionada a la funcionalidad de “Selección de algoritmo para entrenamiento”. (Elaboración propia)

Tabla 12 Historia de usuario: Entrenamiento del modelo seleccionado

REF: HU03	Actor: Usuario
-----------	----------------

Nombre de la Historia	Entrenamiento del modelo seleccionado
Prioridad	Alta
Tiempo estimado	8 horas
Responsable	Alejandro Fidel Lara Miranda
Descripción	Como usuario, quiero poder entrenar el modelo utilizando el algoritmo seleccionado, para preparar el sistema para evaluar imágenes de cáncer de piel.
Criterios de aceptación	El botón "Entrenar" inicia el entrenamiento del modelo con el algoritmo seleccionado.

Nota: Esta tabla describe la historia de usuario relacionada a la funcionalidad de "Entrenamiento del modelo seleccionado". (Elaboración propia)

Tabla 13 Historia de usuario: Evaluación de imágenes de cáncer de piel

REF: HU04	Actor: Usuario
Nombre de la Historia	Evaluación de imágenes de cáncer de piel
Prioridad	Alta
Tiempo estimado	4 horas
Responsable	Stefania Hernández Quiroz
Descripción	Como usuario, quiero poder cargar una imagen de la piel y evaluar si es un melanoma utilizando el algoritmo seleccionado, para recibir un diagnóstico preliminar.
Criterios de aceptación	El sistema muestra los resultados, incluyendo la clase predicha y el porcentaje de confianza en la predicción.

Nota: Esta tabla describe la historia de usuario relacionada a la funcionalidad de "Evaluación de imágenes de cáncer de piel". (Elaboración propia)

3.1.4 Base de datos

3.1.4.1 Descripción general

La base de datos del proyecto de detección de cáncer de piel se encuentra almacenada localmente en la carpeta del proyecto. Esta base de datos contiene un conjunto de imágenes utilizadas para entrenar y evaluar modelos de aprendizaje automático, con el objetivo de detectar diferentes tipos de cáncer de piel. A continuación, se detallan las características y la distribución de las imágenes almacenadas en la base de datos.

:

3.1.4.2 Datos del conjunto de imágenes

La base de datos del proyecto de detección de cáncer de piel se encuentra almacenada localmente en la carpeta del proyecto. Esta base de datos contiene un conjunto de imágenes utilizadas para entrenar y evaluar modelos de aprendizaje automático, con el objetivo de detectar diferentes tipos de cáncer de piel. A continuación, se detallan las características y la distribución de las imágenes almacenadas en la base de datos. El conjunto de datos está compuesto por un total de 6,341 imágenes, que se han clasificado en varias categorías de cáncer de piel. Las imágenes se han recopilado de diversas fuentes y se han almacenado en la carpeta del proyecto para su uso en el sistema de detección. La distribución de imágenes por categoría es la siguiente:

Tabla 14 Cantidad de imágenes por cada tipo de cáncer

TIPO DE CÁNCER	CANTIDAD DE IMÁGENES
Actinic Keratosis	1,356 imágenes
Basal Cell Carcinoma	1,426 imágenes
Dermatofibroma	391 imágenes
Melanoma	4,078 imágenes
Nevus	907 imágenes
Pigmented Benign Keratosis	1,323 imágenes
Seborrheic Keratosis	1,317 imágenes
Squamous Cell Carcinoma	1,435 imágenes
Vascular Lesion	1,435 imágenes

3.1.4.3 Clasificación de imágenes

En la base de datos, Melanoma es la categoría principal con 4,078 imágenes, lo que indica una mayor disponibilidad de datos para esta clase en comparación con las otras categorías. Las demás categorías tienen una cantidad menor de imágenes, que se utiliza para proporcionar una representación equilibrada y diversa del conjunto de datos.

3.1.4.4 Estructura de almacenamiento

Las imágenes están organizadas en carpetas dentro de la estructura del proyecto para facilitar su acceso y manejo. Cada categoría de cáncer de piel tiene una carpeta correspondiente que contiene las imágenes clasificadas. El almacenamiento local permite un acceso rápido y directo a los datos para el entrenamiento y evaluación de los modelos de aprendizaje automático.

3.1.4.5 Descripción de la base de datos ISIC

.La base de datos utilizada en el proyecto proviene de la International Skin Imaging Collaboration (ISIC), una organización reconocida que proporciona un extenso conjunto de imágenes dermatológicas para la investigación y el desarrollo en el área de la detección de cáncer de piel. ISIC se dedica a la recopilación y estandarización de imágenes de alta calidad que facilitan el avance de los métodos de diagnóstico y la mejora de los algoritmos de aprendizaje automático en el campo de la dermatología.

3.1.4.6 Origen y acceso a la base de datos

La base de datos ISIC es accesible a través del portal de ISIC, que ofrece un repositorio de imágenes dermatológicas utilizadas para el entrenamiento y evaluación de modelos de detección de cáncer de piel. Los datos se recopilan de diversas fuentes clínicas y de investigación, asegurando una amplia representación de condiciones dermatológicas.

Para obtener la base de datos ISIC, se siguió el siguiente proceso:

- Registro en el Portal: Se creó una cuenta en el portal de ISIC para acceder a los datos y cumplir con los requisitos de uso de la base de datos.
- Selección del Conjunto de Datos: Se seleccionaron las imágenes relevantes para el proyecto, enfocándose en las categorías de cáncer de piel específicas que eran necesarias para el entrenamiento del modelo.
- Descarga: Las imágenes fueron descargadas en formato digital desde el portal, asegurando que los datos estuvieran organizados y etiquetados de acuerdo con las categorías definidas por ISIC.

3.1.4.7 Relevancia para el proyecto

.La base de datos ISIC es crucial para el proyecto de detección de cáncer de piel debido a la calidad y precisión de las imágenes proporcionadas. El enfoque en melanoma y otras categorías de cáncer de piel permite entrenar modelos de aprendizaje automático con datos relevantes y bien etiquetados, lo que mejora la capacidad del sistema para realizar diagnósticos precisos.

3.1.5 Desarrollo del aplicativo

3.1.5.1 Diseño de interfaces de usuario

En la primera fase del desarrollo del aplicativo de detección de cáncer de piel, se ha llevado a cabo la creación de mockups utilizando la herramienta de diseño de interfaces Figma. Esta herramienta es esencial para la planificación visual del proyecto, ya que facilita la elaboración de diseños detallados que son cruciales para la estructuración y desarrollo del aplicativo.

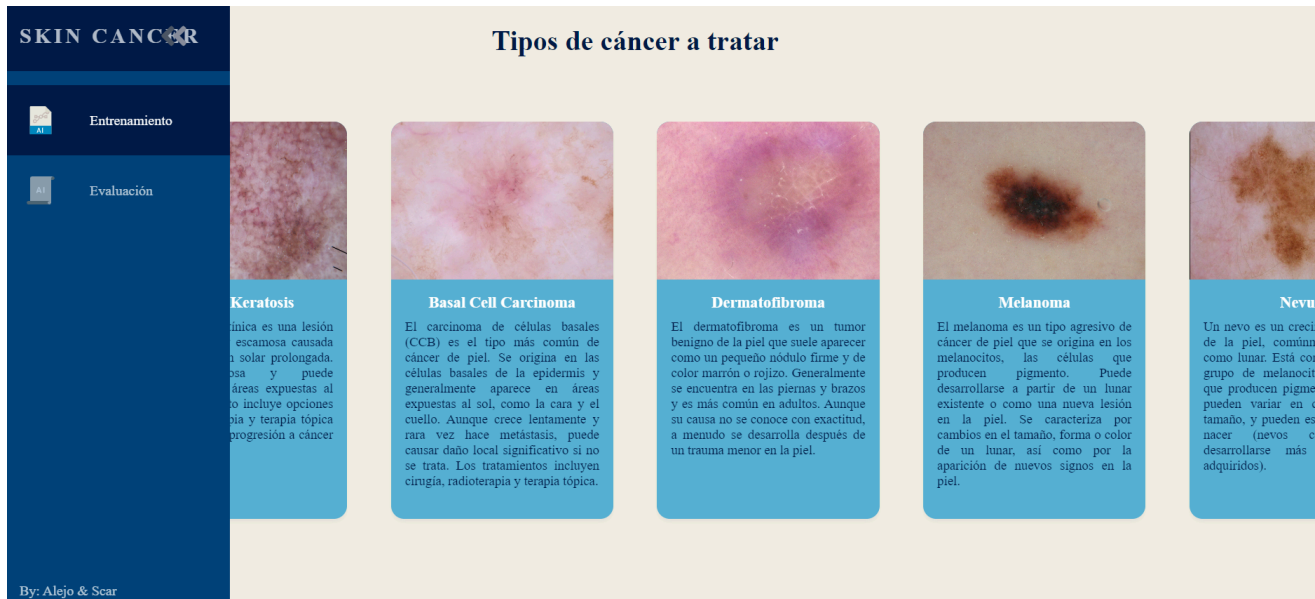
Figma ofrece un entorno de diseño colaborativo que permite la creación de mockups precisos, representando de manera clara la interfaz de usuario y la disposición de las funcionalidades del sistema. Los mockups desarrollados en esta fase cumplen dos objetivos principales. Primero, definen la estética del aplicativo, incluyendo la apariencia visual, la paleta de colores, la tipografía y los elementos gráficos. Esta visualización preliminar asegura que el diseño sea atractivo y coherente con los objetivos del proyecto.

Ilustración 17 Interfaz de navegación inicial



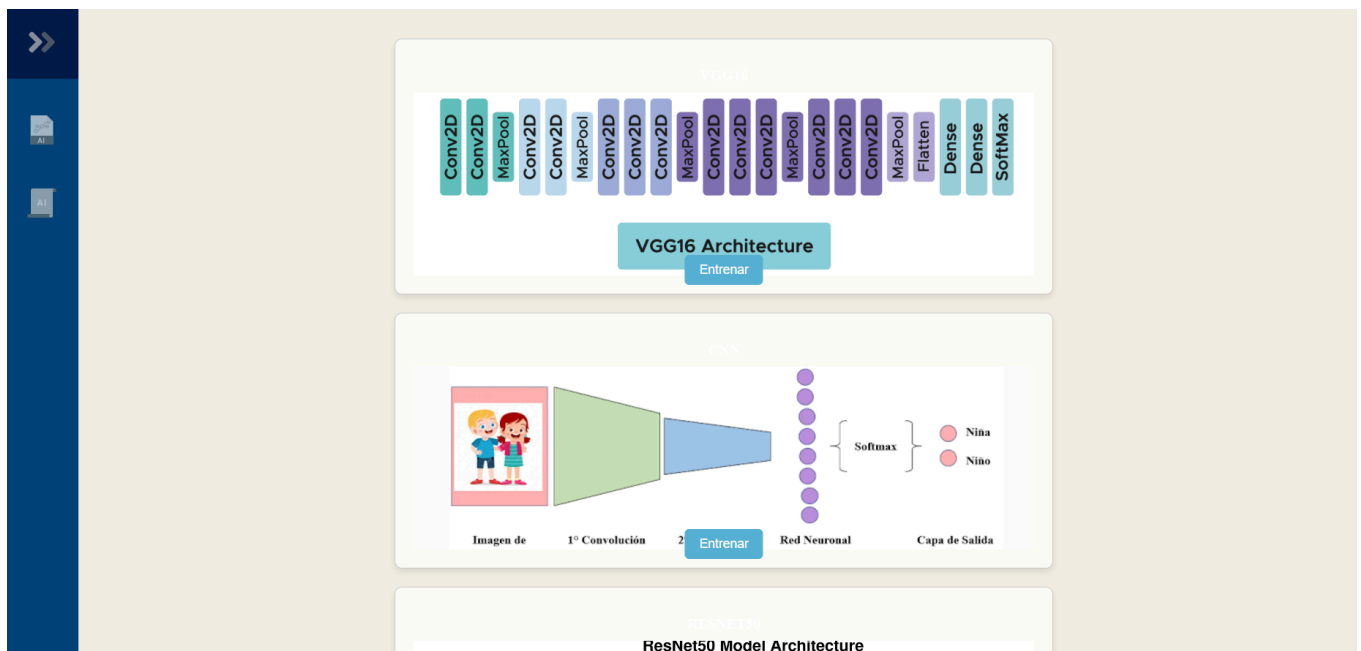
Nota: En esta figura se muestra el inicio donde se explican diferentes tipos de cáncer.

Ilustración 18 Menú desplegable con las opciones a tomar



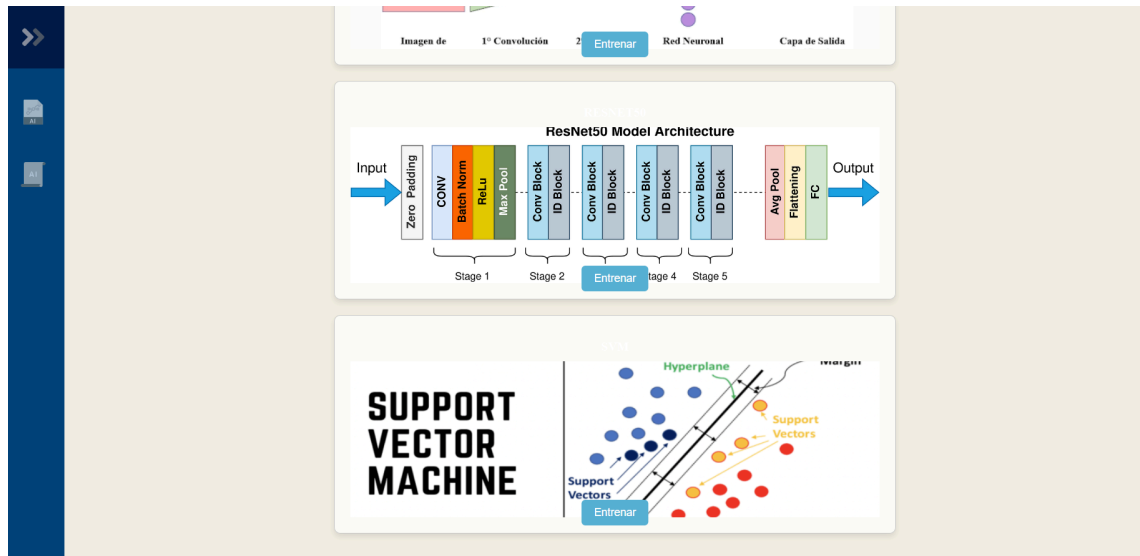
Nota: En esta figura se ve la barra desplegable en la cual se pueden entrar a los otros módulos.

Ilustración 19 Interfaz de entrenamiento



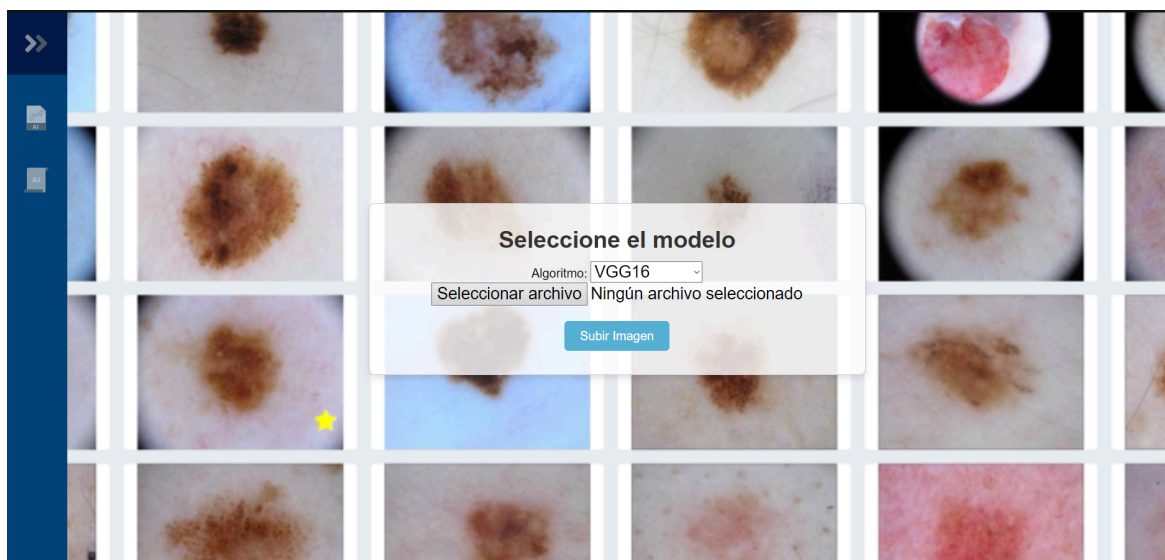
Nota: En esta figura se ven los diferentes tipos de modelos los cuales pueden ser entrenados.

Ilustración 20 Interfaz de entrenamiento



Nota: En esta figura se ve la continuación de los diferentes tipos de modelos los cuales pueden ser entrenados.

Ilustración 21 Interfaz de evaluación



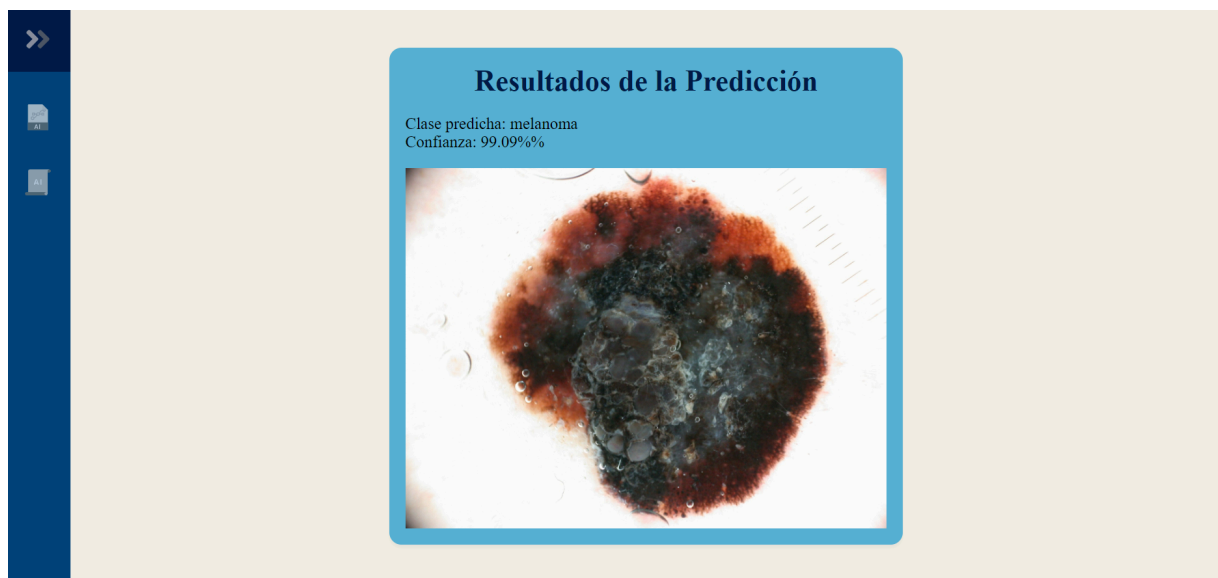
Nota: En esta figura se ve la interfaz de evaluación según el modelo a escoger.

Ilustración 22 Interfaz de selección de evaluación



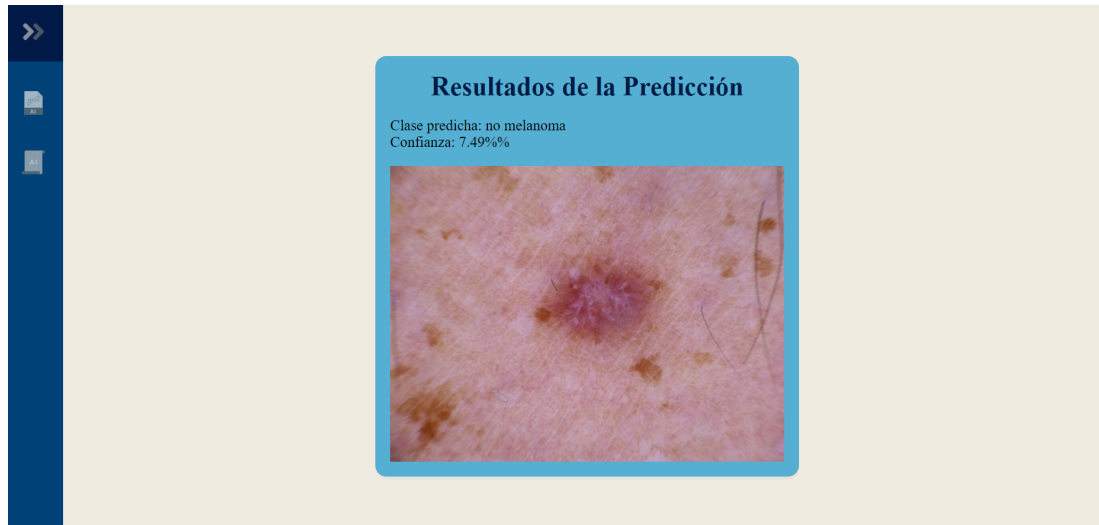
Nota: En esta figura se ve la selección de los diferentes modelos a entrenar

Ilustración 23 Interfaz de resultado positivo de la predicción



Nota: En esta figura se ve la interfaz de evaluación según el modelo a escoger.

Ilustración 24 Interfaz de resultado negativo de la predicción



Nota: En esta figura se ve la interfaz de evaluación según el modelo a escoger.

3.1.5.2 Desarrollo del backend

El desarrollo del backend de la aplicación se realizó utilizando Python con el microframework Flask, seleccionado por su ligereza y flexibilidad para el desarrollo eficiente de aplicaciones web. A continuación, se detalla cómo se configuró y desarrolló este backend, incluyendo la integración con el modelo de aprendizaje profundo, la configuración del entorno de desarrollo, y las principales funcionalidades implementadas.

Configuración del Entorno

Para el desarrollo del backend, se utilizó Visual Studio Code como el entorno de desarrollo integrado (IDE) principal, complementado con entornos virtuales para gestionar las dependencias de manera aislada y eficiente. El desarrollo se llevó a cabo en un sistema operativo Windows, aprovechando la compatibilidad con las herramientas necesarias para el proyecto. En lugar de utilizar un proveedor en la nube como AWS, el despliegue y las pruebas de la aplicación se realizaron de manera local. Esto permitió un control total sobre el entorno de desarrollo y simplificó la gestión de recursos durante la fase inicial del proyecto.

Herramientas principales de desarrollo

- Visual Studio Code: IDE principal utilizado para el desarrollo.
- Lenguaje de Programación: Python 3.10.

- Framework: Flask.
- Entornos Virtuales: Uso de virtualenv para aislar dependencias.
- Sistema Operativo: Windows, seleccionado por su compatibilidad y flexibilidad.

-
- Modelado y Entrenamiento de Red Neuronal: Utilización de TensorFlow y Keras para desarrollar el modelo de detección de melanoma.
 - Almacenamiento Local: La base de datos y las imágenes se gestionan localmente dentro del sistema de archivos.

Desarrollo del modelo de detección de melanoma

Se integró un modelo de aprendizaje profundo utilizando DenseNet121 preentrenado, adaptado para la detección de melanomas. Las imágenes se procesan y se etiquetan según su clasificación, donde las imágenes de melanomas se etiquetan con 1, y el resto con 0. El modelo se entrena utilizando un generador de datos personalizado que ajusta dinámicamente las imágenes para el entrenamiento.

El modelo final se guarda en el sistema de archivos local en formato Keras y también se almacenan los pesos entrenados para futuras evaluaciones. Este enfoque asegura que el modelo esté optimizado para el entorno específico del proyecto.

Funcionalidades Principales

- Carga y procesamiento de imágenes: El backend permite la carga de imágenes y su procesamiento utilizando el modelo entrenado.
- Predicción de melanoma: El modelo clasifica las imágenes como melanoma o no melanoma, mostrando los resultados al usuario.
- Interfaz de usuario: A través de Flask, se desarrolló una API sencilla que permite interactuar con el modelo desde el frontend, facilitando la integración con otras partes del sistema.

3.1.5.2.1 Integraciones

1. **Flask** fue el microframework principal utilizado para desarrollar el backend de la aplicación. Su simplicidad y flexibilidad permitieron una rápida implementación de la lógica de negocio y la exposición de endpoints para interactuar con el frontend.

Uso: Flask se utilizó para manejar las rutas HTTP, procesar las solicitudes y responder con datos, incluyendo los resultados de las predicciones realizadas por el modelo de detección de cáncer de piel.

2. **TensorFlow, junto con Keras**, se utilizó para implementar y entrenar el modelo de aprendizaje profundo basado en DenseNet121 para la detección de melanoma.

Uso: Las capas de procesamiento de imágenes y el modelo fueron configurados y

entrenados utilizando estas bibliotecas. Además, TensorFlow se integró en el backend para cargar el modelo preentrenado y realizar predicciones en tiempo real.

3.1.5.3 Desarrollo del frontend

El desarrollo del frontend para el proyecto de detección de cáncer de piel se llevó a cabo en conjunto con el backend desarrollado en Python y Flask. Esta combinación permitió la creación de una aplicación web robusta, donde el frontend interactúa de manera fluida con el backend para ofrecer a los usuarios una experiencia de uso eficiente y sin complicaciones.

El frontend fue diseñado utilizando HTML5 y CSS3, lo que proporcionó una estructura clara y un diseño atractivo para las páginas web. Estas tecnologías permitieron crear una interfaz que es tanto visualmente agradable como intuitiva, facilitando la navegación para los usuarios. Además, se prestó especial atención a la usabilidad, asegurando que los usuarios pudieran interactuar con la aplicación de manera sencilla, incluso sin conocimientos técnicos avanzados.

Para añadir interactividad y mejorar la experiencia del usuario, se empleó JavaScript junto con jQuery. Estas herramientas fueron fundamentales para implementar funcionalidades dinámicas, como la validación en tiempo real de los formularios de carga de imágenes. De esta manera, los usuarios pueden subir imágenes de manera rápida y recibir retroalimentación instantánea, todo sin necesidad de recargar la página.

El uso de Flask permitió integrar perfectamente el frontend con el backend. A través de Flask, las solicitudes del usuario, como la carga de imágenes para su análisis, son procesadas de manera eficiente por el servidor. Flask también facilitó la creación de rutas y la gestión de la comunicación entre el frontend y el backend, asegurando que los datos se transmitan y se procesen de forma segura y rápida.

Finalmente, los mockups del frontend fueron diseñados en Figma antes de comenzar con la codificación. Estos diseños preliminares sirvieron para planificar y visualizar la estructura de la interfaz de usuario, permitiendo que el equipo de desarrollo alineara las funcionalidades del sistema con las necesidades específicas de los usuarios. El resultado es una interfaz de usuario intuitiva, funcional y estéticamente agradable, que facilita la interacción con el sistema y contribuye a la eficacia del proyecto en la detección de cáncer de piel.

3.2 – ANÁLISIS DE RESULTADOS Y DISCUSIÓN

3.2.1 Relación con el objetivo general

El objetivo general del proyecto fue desarrollar un sistema para la detección de cáncer de piel utilizando técnicas de aprendizaje automático. Los resultados obtenidos muestran que el sistema desarrollado cumple con este objetivo al proporcionar una herramienta efectiva para la clasificación de imágenes dermatológicas. La implementación del modelo de aprendizaje automático ha permitido identificar y clasificar diferentes tipos de cáncer de piel con un alto grado de precisión. La capacidad del sistema para procesar imágenes y generar resultados de clasificación en tiempo real demuestra una alineación efectiva con el objetivo general, ofreciendo una solución valiosa para el diagnóstico temprano del cáncer de piel.

3.2.2 Evaluación respecto a los objetivos específicos

Selección del banco de datos de imágenes: Se seleccionó el banco de datos ISIC (International Skin Imaging Collaboration), que es ampliamente reconocido por su calidad y diversidad de imágenes dermatológicas. La elección de este banco de datos se basó en su exhaustiva colección de imágenes etiquetadas, que facilita el entrenamiento y la evaluación del modelo de detección de cáncer de piel.

Preprocesamiento y procesamiento de imágenes: Se llevó a cabo un exhaustivo proceso de preprocesamiento, que incluyó la normalización de las imágenes, el redimensionamiento y la conversión a formatos adecuados para el entrenamiento del modelo. Este procesamiento inicial fue esencial para asegurar que las imágenes fueran consistentes y adecuadas para el análisis por parte del modelo de aprendizaje automático.

Evaluación de diferentes arquitecturas de redes neuronales: Se evaluaron varias arquitecturas de redes neuronales, incluyendo VGG16, ResNet50 y DenseNet121. Cada modelo fue probado con el banco de datos seleccionado para determinar su eficacia en la clasificación de imágenes de cáncer de piel. La evaluación comparativa permitió identificar la arquitectura más adecuada para el problema específico de detección.

Entrenamiento de la red neuronal seleccionada: El entrenamiento de la red neuronal se realizó utilizando el banco de datos de imágenes procesadas. Se implementaron técnicas de transferencia de aprendizaje y ajuste fino para mejorar el rendimiento del modelo. Los resultados del entrenamiento mostraron una mejora en la precisión y la capacidad de generalización del modelo sobre nuevas imágenes.

Desarrollo de módulos para carga, procesamiento, entrenamiento y generación de resultados: Se desarrollaron y probaron diversos módulos del sistema, incluyendo la carga de imágenes, el procesamiento de datos, el entrenamiento del modelo, y la generación de resultados. Estos módulos se integraron en una plataforma cohesiva que permite al usuario cargar imágenes, procesarlas, y recibir predicciones sobre la presencia de cáncer de piel.

Implementación del sistema web para la detección de cáncer de piel: Se construyó una aplicación web utilizando Flask para proporcionar una interfaz amigable para el usuario. La aplicación permite la carga de imágenes, el procesamiento a través del modelo de aprendizaje automático, y la visualización de los resultados de detección. La implementación de esta herramienta web facilita el acceso a la tecnología de detección de cáncer de piel para usuarios finales y profesionales de la salud.

3.2.3 Discusión sobre la metodología aplicada

La metodología ágil Kanban fue adoptada para el desarrollo del proyecto, proporcionando una estructura flexible y adaptable para gestionar el progreso y las tareas del proyecto. Esta metodología permitió una organización eficiente de las actividades, facilitando la priorización de tareas y la respuesta rápida a cualquier desafío que surgiera durante el desarrollo. La implementación de Kanban resultó ser efectiva para mantener el proyecto en curso, optimizando el flujo de trabajo y asegurando que se cumplieran los objetivos del proyecto en tiempo y forma. La combinación de técnicas avanzadas de aprendizaje automático con una metodología de gestión ágil ha permitido la creación de un sistema robusto y eficaz para la detección de cáncer de piel.

3.3 – CONCLUSIONES

El desarrollo del sistema de detección de cáncer de piel utilizando técnicas de aprendizaje automático ha alcanzado los objetivos propuestos y ha demostrado ser una herramienta efectiva en el diagnóstico precoz de esta enfermedad. A continuación se presentan las conclusiones clave del proyecto:

Desempeño del sistema: El sistema desarrollado ha logrado implementar un modelo de aprendizaje automático robusto para la detección de cáncer de piel, alcanzando altos niveles de precisión en la clasificación de imágenes dermatológicas. La selección y el preprocesamiento del banco de datos ISIC han sido fundamentales para asegurar la calidad y la efectividad del modelo.

Interfaz de usuario eficiente: La implementación de una interfaz web utilizando Flask ha facilitado la interacción del usuario con el sistema, permitiendo la carga, procesamiento y análisis de imágenes de manera sencilla y efectiva. La experiencia de usuario ha sido optimizada para garantizar que el proceso de detección sea accesible y comprensible para usuarios no técnicos.

Evaluación de modelos: La evaluación de diferentes arquitecturas de redes neuronales, como VGG16, ResNet50 y DenseNet121, ha permitido identificar el modelo más adecuado para la tarea de detección de cáncer de piel. El entrenamiento y ajuste fino de estos modelos han resultado en un sistema que ofrece resultados de alta confianza y precisión.

Manejo de datos y entrenamiento: El procesamiento de imágenes y la implementación de técnicas de transferencia de aprendizaje han sido esenciales para mejorar la capacidad de generalización del modelo. El uso de una base de datos bien estructurada y el desarrollo de módulos específicos para cada etapa del proceso han contribuido al éxito del proyecto.

Implementación y escalabilidad: La elección de tecnologías y herramientas, como Python y Flask, ha permitido desarrollar un sistema escalable y sostenible. La metodología ágil Kanban aplicada en el desarrollo del proyecto ha facilitado una gestión eficiente de las tareas y ha permitido ajustes continuos basados en los resultados y el feedback obtenido.

Contribución al diagnóstico: El sistema no solo cumple con los objetivos técnicos del proyecto, sino que también representa una contribución significativa al diagnóstico temprano del cáncer de piel. La capacidad de analizar imágenes de manera precisa y en tiempo real ofrece un recurso valioso para la detección y tratamiento oportuno de esta enfermedad.

3.4 – RECOMENDACIONES

Para optimizar aún más el sistema de detección de cáncer de piel y asegurar su eficacia continua, se recomienda considerar las siguientes acciones:

Primero, ampliar el conjunto de datos con más imágenes variadas, especialmente aquellas

que incluyan casos raros y menos representados, como tipos infrecuentes de cáncer y lesiones benignas. Esta ampliación contribuirá a mejorar la capacidad del modelo para generalizar y reducir el riesgo de sesgo, asegurando una detección más precisa en escenarios clínicos diversos.

En segundo lugar, implementar técnicas de aumento de datos para enriquecer el entrenamiento del modelo. Las técnicas de aumento, como rotaciones, cambios en el brillo y escalado, pueden ayudar a simular variaciones en las imágenes y a hacer el modelo más robusto frente a diferentes condiciones de iluminación y ángulos.

Además, realizar una validación continua y pruebas de rendimiento del sistema en entornos clínicos reales es crucial para evaluar su precisión y confiabilidad. Las pruebas regulares con datos nuevos y casos reales permitirán identificar áreas de mejora y ajustar el modelo para mantener altos niveles de rendimiento y exactitud.

También se recomienda integrar el sistema con herramientas de análisis avanzadas que permitan un análisis más profundo y detallado de los resultados, como la interpretación de características específicas en las imágenes y la correlación con diagnósticos clínicos.

Finalmente, es fundamental fomentar la colaboración con profesionales de la salud para obtener retroalimentación valiosa sobre la funcionalidad del sistema y su utilidad en el entorno clínico. La colaboración continua garantizará que el sistema se adapte a las necesidades cambiantes y proporcione un apoyo efectivo a los médicos en la detección temprana del cáncer de piel.

3.5 BIBLIOGRAFÍA

- [1] Fuentes, A. S. F., Cabrera, R. G., & Rodriguez, E. V. (2022). Identificación automática de cáncer de piel aplicando Machine Learning. REVISTA COLOMBIANA DE TECNOLOGIAS DE AVANZADA (RCTA), 2(40).
- [2] Pozzobon, F., Acosta, Á., & Castillo, J. (2018). Cáncer de piel en Colombia: cifras del Instituto Nacional de Cancerología. *asocolderma*, 26(1). doi:<https://doi.org/10.29176/2590843X.25>
- [3] THH Aldhyani, M. Alrasheed, MH Al-Adaileh, AA Alqarni, MY Alzahrani y AH Alahmadi, “Aprendizaje profundo y algoritmos de tendencia holt para predecir la pandemia de covid-19”. *Computadoras, Materiales y Continua*, vol. 67, núm. 2, págs. 2141-2160, 2021.
- [4] Alsaade FW, Aldhyani THH, AlAdhaileh MH. Developing a Recognition System for Diagnosing Melanoma Skin Lesions Using Artificial Intelligence Algorithms. *Comput Math Methods Med*. 2021 May 15;2021:9998379. doi: 10.1155/2021/9998379. PMID: 34055044; PMCID: PMC8143893.
- [5] Curci Robledo, G. J., Méndez Garabetti, M., & Sandez, P. J. (2023). Inteligencia artificial: herramienta diagnóstica para el cáncer de mama. In XXVIII Congreso Argentino de Ciencias de la Computación (CACIC)(La Rioja, 3 al 6 de octubre de 2022). <http://sedici.unlp.edu.ar/handle/10915/149645>
- [6] Guzman-Loreto, G. R., Pérez-Diosdado, A. Z., Durán-Reyes, E., Torres-Calderon, A., González-Guzman, L., Preciado-Trujillo, H., & Guzman-Cabrera, R. (2023). Aplicación de la inteligencia artificial en la clasificación automática de imágenes médicas. *JÓVENES EN LA CIENCIA*, 21, 1–6. Recuperado a partir de <https://www.jovenesenlaciencia.ugto.mx/index.php/jovenesenlaciencia/article/view/4038>
- [7] Vega, M. Á., Mora, L. M. Q., & Badilla, M. V. C. (2020). Inteligencia artificial y aprendizaje automático en medicina. *Revista médica sinergia*, 5(8), e557-e557.
- [8] Moreno Villalba, L. M., Ávila Camacho, F. J., Meléndez Ramírez, A., & Manuel Stein, J. (2017). Modelo para la detección de melanoma mediante el escaneo de imágenes con Deep Learning usando el algoritmo K-NN(Abril 2018). *TecnoCultura*, 8. Recuperado a partir de <https://tecnocultura.org/index.php/Tecnocultura/article/view/230>
- [9] Jaime-Calva, L. R., Castillejos-Fernández, H., Franco-Árcega, A., Miranda-Romagnoli, P., &

- Pérez-Cortés, O. (2022). Clasificación de cáncer de piel utilizando aprendizaje profundo. *Pä*
Boletín Científico De Ciencias Básicas E Ingenierías Del ICBI, 10(Especial3), 147-152.
<https://doi.org/10.29057/icbi.v10iEspecial3.9029>
- [10] Ley 1581 de 2012 - Gestor Normativo - Función Pública. Recuperado de:
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981>.
-
- [11] Decreto 1377 de 2013 - Gestor Normativo - Función Pública. Recuperado de:
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=53646>.
- [12] Ley 1341 de 2009 - Gestor Normativo. (s. f.). Función Pública.
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=36913>
- [13] Política Nacional para la Transformación Digital e Inteligencia Artificial | SITEAL. (s. f.).
<https://siteal.iiep.unesco.org/bdnp/3166/politica-nacional-transformacion-digital-inteligencia-artificial>
- [14] Sharanharsoor. (2023, 16 marzo). Skin cancer detection. Kaggle.
<https://www.kaggle.com/code/sharanharsoor/skin-cancer-detection/notebook>
- [15] López Martínez, Guillermo (2020). Segmentación de lesiones de piel en imágenes dermoscópicas mediante redes neuronales profundas. Proyecto Fin de Carrera / Trabajo Fin de Grado, E.T.S.I. y Sistemas de Telecomunicación (UPM), Madrid.
- [16] Márquez Díaz, Jairo. (2020). Visión artificial profunda aplicada a la identificación temprana de cáncer no melanoma y queratosis actínica Deep Artificial Vision Applied to the Early Identification of Non-Melanoma Cancer and Actinic Keratosis. *Computacion y Sistemas*. 24. 751-766.
10.13053/CyS-24-2-2901.
- [17] Alcoceba Álvarez, D. (2018). Detector automático de lesiones de piel en imágenes dermoscópicas con deep learning.
- [18] Hekler, A., Utikal, J., Enk, A., Hauschild, A., Weichenthal, M., Maron, R. C., Berking, C., Haferkamp, S., Klode, J., Schadendorf, D., Schilling, B., Holland-Letz, T., Izar, B., Von Kalle, C., Fröhling, S., Brinker, T. J., Schmitt, L., Peitsch, W. K., Hoffmann, F., . . . Thiem, A. (2019). Superior skin cancer classification by the combination of human and artificial intelligence. *European Journal of Cancer*, 120, 114-121. <https://doi.org/10.1016/j.ejca.2019.07.019>

Universidad Popular del Cesar
Programa de Ingeniería de Sistemas

- [19] Villalba, L. M. M., Camacho, F. J. Á., Ramírez, A. M., & Stein, J. M. (2017). Modelo para la detección de melanoma mediante el escaneo de imágenes con Deep Learning usando el algoritmo K-NN (Abril 2018). TecnoCultura, 8-8.

Universidad Popular del Cesar
Programa de Ingeniería de Sistemas



Valledupar, 10 de septiembre del 2024.

Señores:

COMITÉ DE INVESTIGACIÓN
Facultad de Ingenierías y Tecnológicas
Programa de Ingeniería de Sistemas
Universidad Popular Del Cesar

Cordial saludo

Yo **BRAULIO BARRIOS ZUÑIGA**, identificado con la cédula de ciudadanía No. **73107112**, certifico que he revisado el documento correspondiente al proyecto que lleva por título "**SISTEMA DE APRENDIZAJE AUTOMÁTICO PARA LA DETECCIÓN DE CÁNCER DE PIEL**", presentada por los estudiantes **STEFANIA HERNANDEZ QUIROZ** y **ALEJANDRO FIDEL LARA MIRANDA**, y, después de haberle realizado las respectivas correcciones, cuenta con mi aprobación para ser presentada ante el comité. Sugiero la aprobación por parte de ustedes.

Línea de investigación: Transformación digital

Sublínea: Big Data y Analytics

Agradezco la atención prestada

Atentamente,

BRAULIO BARRIOS ZUÑIGA
CC 73107112 de Cartagena de Indias
Director del Proyecto

Universidad Popular del Cesar
Programa de Ingeniería de Sistemas



Anexo B. Modelo Carta de los estudiantes

Valledupar 10 de septiembre del 2024

Señores:

COMITÉ DE INVESTIGACIÓN

Facultad de Ingenierías y Tecnológicas

Programa de Ingeniería de Sistemas

Universidad Popular Del Cesar

Cordial saludo

Nosotros los abajo firmantes, estudiantes del programa de Ingeniería de sistemas, presentamos a ustedes el documento correspondiente al proyecto de grado denominado **“SISTEMA DE APRENDIZAJE AUTOMATICO”**.

Quedamos a la espera del concepto emitido por el comité respecto de la viabilidad y aceptación de dicha propuesta.

Agradecemos la atención prestada

Atentamente,

STEFANIA HERNANDEZ QUIROZ

CC. 1003377924 de Valledupar

ALEJANDRO FIDEL LARA MIRANDA

CC. 1004282874 de Valledupar

Universidad Popular del Cesar
Programa de Ingeniería de Sistemas



Universidad Popular del Cesar
Programa de Ingeniería de Sistemas



Anexo C . Carta modelo aval de la entidad responsable

Valledupar, 10 de septiembre del 2024.

Señores:

COMITÉ DE INVESTIGACIÓN

Facultad de Ingenierías y Tecnológicas

Programa de Ingeniería de Sistemas

Universidad Popular Del Cesar

Cordial Saludo respetados Ingenieros,

Por medio de la presente, deseo comunicar que como estudiante de la Universidad Popular Del Cesar, me encuentro en proceso de elaboración de mi trabajo de grado titulado "Sistema de aprendizaje automático para la detección de cáncer de piel". Es importante resaltar que no cuento con el respaldo de ninguna empresa o entidad externa para este proyecto académico.

Este proyecto se llevará a cabo de manera independiente, utilizando una base de datos pública disponible en Kaggle (referencia: [<https://www.kaggle.com/code/sharanharsoor/skin-cancer-detection/input>]). Dada la naturaleza abierta del proyecto y la ausencia de una entidad específica, el sistema se centrará en la creación de un prototipo que será evaluado exhaustivamente utilizando la base de datos mencionada. Estos repositorios son reconocidos por la calidad y fiabilidad de la información que albergan, y su uso se ajusta a los lineamientos éticos y académicos establecidos por la institución.

Atentamente,

Stefania Scarlet Hernández Quiroz

CC. 1.003.377.924 de Valledupar

Alejandro Fidel Lara Miranda

CC. 1.004.282.874 de Valledupar

Anexo D. Carta que evidencia la entrega del artículo científico

Machine learning system for skin cancer detection

Sistema de aprendizaje automático para la detección de cáncer de piel

Stefania Scarlet Hernández Quiroz

sscarlethernandez@unicesar.edu.co

Alejandro Fidel Lara Miranda

afidellara@unicesar.edu.co

RESUMEN

Este artículo presenta un sistema de detección de melanoma utilizando el modelo de red neuronal convolucional DenseNet121, implementado en Python. El sistema se integra con una interfaz web desarrollada en Flask para permitir la carga y clasificación de imágenes de piel. El proceso incluye la preparación de datos, el entrenamiento del modelo, y la integración del modelo entrenado en una aplicación web para proporcionar una herramienta accesible para la detección temprana de cáncer de piel.

ABSTRACT

This paper presents a melanoma detection system using the DenseNet121 convolutional neural network model, implemented in Python. The system integrates with a web interface developed in Flask to allow the upload and classification of skin images. The process includes data preparation, model training, and the integration of the trained model into a web application to provide an accessible tool for early skin cancer detection.

INTRODUCCIÓN

La detección temprana del melanoma, un tipo de cáncer de piel, es crucial para mejorar las tasas de supervivencia. Este artículo presenta un sistema basado en aprendizaje profundo utilizando

DenseNet121 para clasificar imágenes de piel en melanoma y no melanoma. El sistema está diseñado para ser accesible a través de una interfaz web creada con Flask, permitiendo a los usuarios subir imágenes y obtener predicciones en tiempo real.

METODOLOGÍA

Preparación de datos

El sistema utiliza el conjunto de datos "Skin cancer ISIC The International Skin Imaging Collaboration", que contiene imágenes clasificadas en diferentes tipos de cáncer de piel. Para este estudio, se clasifican las imágenes en dos categorías: melanoma (1) y no melanoma (0). El proceso de preparación de datos incluye:

- Generación de Datos: Utilización de ImageDataGenerator para la normalización de imágenes.
- Asignación de Etiquetas: Implementación de una función personalizada para etiquetar las imágenes como melanoma o no melanoma basándose en la estructura de directorios.

Generador de datos personalizado

Se desarrolla un generador de datos personalizado (CustomDataGenerator) que carga imágenes y etiquetas de manera eficiente en lotes para el entrenamiento del modelo. Este generador permite la adaptación del tamaño de las imágenes y el

Universidad Popular del Cesar
Programa de Ingeniería de Sistemas