

## Proyecto 1 - Tópicos avanzados en analítica

### **Detección de Frailejones (también conocida como ‘Big Monk’ Detection) sobre imágenes aéreas del páramo de Chingaza y Cruz Verde usando la metodología CRISP-DM**

*Luis Alejandro Garzón Ramirez*

*Zorayda Acevedo Jimenez*

*Yulieth Danitza Aguillón Ortega*

#### **1. Entendimiento de la situación**

El agua es uno de los recursos naturales más esenciales, ya que es indispensable para la supervivencia de todos los seres vivos. De acuerdo con el Instituto Humboldt (2022), entidad colombiana, vinculada al Ministerio de Ambiente y Desarrollo Sostenible, regida por el derecho privado, que investiga acerca de la biodiversidad y de las relaciones entre esta y el bienestar humano menciona que el frailejón (Subtribu Espeletiinae) es una de las especies emblemáticas de los páramos del país, ya que el frailejón juega un papel esencial en la regulación de los ecosistemas de páramo. Gracias a sus hojas que capturan la humedad, esta planta ayuda a mantener el balance hídrico en la región, regulando el flujo de agua hacia los ríos y manteniendo la humedad en el suelo, lo que beneficia tanto a la flora como a la fauna local. Colombia cuenta hoy con 92 especies de frailejones. La importancia de los frailejones contrasta con la escasez de investigaciones e información sobre su ecología, en particular acerca de los efectos que diferentes presiones naturales o antrópicas tienen sobre sus ciclos de vida y la dinámica de las poblaciones. El cambio climático es la causa más probable de la afectación de la densidad de los frailejones, una confluencia de agentes bióticos que incluyen insectos herbívoros y hongos, que reducen su crecimiento, aumentan su tasa de mortalidad y disminuyen la capacidad de retención, de almacenamiento y de intercepción de agua. De allí surgió el interés por parte de varias instituciones como: la Sociedad Colombiana de Entomología, la Unidad Administrativa Especial del Sistema de Parques Nacionales Naturales de Colombia y Patrimonio Natural, Fondo para la Biodiversidad y las Áreas Protegidas, para trabajar en identificar las causas de la afectación de los frailejones y así establecer las medidas participativas para su manejo y control.

La caracterización, diagnóstico, evaluación y monitoreo del estado de la riqueza de la flora, permite entre otros beneficios la identificación de aquellas especies que por su valor, importancia o estado de amenaza se constituyen en objeto y prioridad de conservación, a su vez que son indicadoras del estado de los diferentes ecosistemas. A partir de ese conocimiento e información, se puede de manera oportuna generar modelamientos y priorización de áreas de interés como corredores de conservación, áreas de manejo especial y acciones concretas y eficaces para la protección y conservación de esa riqueza natural. (Corporación Autónoma Regional de Cundinamarca-CAR; 2018). Es en este contexto cobra importancia el uso de metodologías avanzadas de análisis de datos para identificar, preservar y monitorear los ecosistemas de páramo, hábitat natural de los frailejones. Dado su alto potencial hídrico, resulta fundamental determinar las zonas donde se encuentran estas especies y aplicar este análisis en otras áreas, facilitando la restauración del ecosistema en regiones con baja densidad de poblaciones de *Espeletia*.

## 2. Objetivos de negocio

2.1 Estimar la Probabilidad de que una planta sea un Frailejón (Espeletia): Desarrollar un modelo de detección que permita predecir con precisión la probabilidad de que una planta observada sea un frailejón optimizando la identificación y clasificación de esta especie.

2.2 Calcular la Densidad de Frailejones en un Área Determinada: Analizar y cuantificar la distribución de frailejones en distintas regiones para evaluar su densidad poblacional, facilitando estudios de conservación y monitoreo ambiental.

## 3. Objetivos de minería de datos

**3.1 Objetivo general:** construir una red neuronal para detectar automáticamente frailejones sobre imágenes aéreas del páramo de Chingaza y Cruz Verde.

### 3.2 Objetivos específicos:

3.2.1. Diseñar seis (6) propuestas de modelos usando redes neuronales ajustando arquitecturas e hiperparámetros.

3.2.2 Obtener un desempeño en los modelos medido por el AUC de la siguiente forma para cada modelo:

- Red Sencilla:  $> 0.88$
- Multi-capas:  $> 0.95$
- Convolutiva:  $> 0.95$
- Convolutiva + VGG16  $> 0.99$
- Convolutiva con transferencia VGG-16 y ajuste de hiperparámetros - Sin AUC objetivo
- Otra red pre-entrenada: MobileNetV2 – Sin AUC objetivo

3.2.3 Verificar de forma visual (cualitativa) la detección de frailejones en el mejor modelo

## 4. Entendimiento de los datos

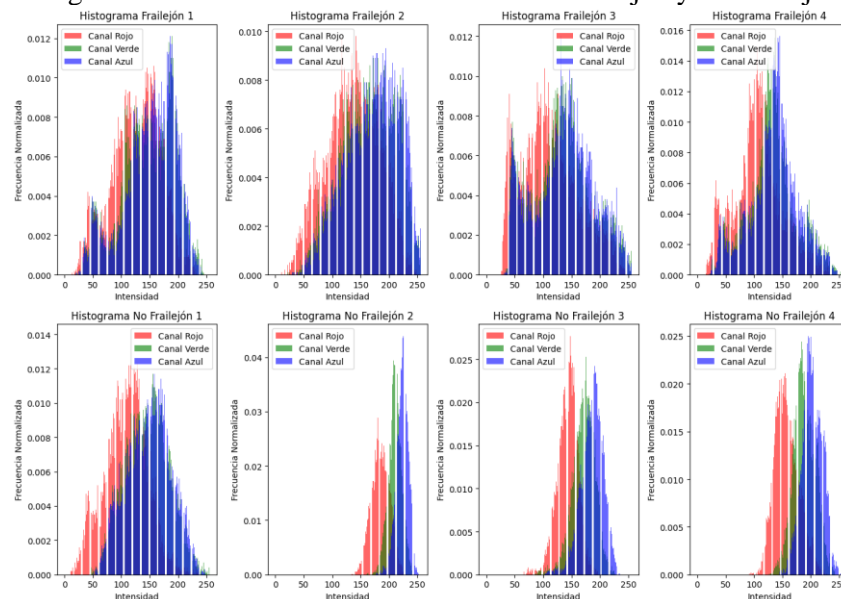
Se dispone de una imagen llamada IMG\_3451 en formato JPG tamaño 3000x4000 pixeles, donde hay presencia de frailejones. Tomamos muestras aleatorias y se visualizan en una cuadrícula para identificar cuando hay presencia o no de frailejones. Se fijan semillas aleatorias para garantizar la reproducibilidad y se seleccionan 4 imágenes aleatorias de cada categoría “Frailejon” y No frailejon”. En la Imagen 1 se observa una figura con subgráficas para mostrar las imágenes seleccionadas. Las imágenes de Frailejón se muestran en la fila superior y las de No Frailejón en la fila inferior. Las imágenes están en formato RGB y poseen el mismo tamaño, lo que las hace adecuadas para su uso en modelos de clasificación o segmentación de imágenes (ver imagen 1).

Imagen 1. Visualización de muestra de datos: Frailejon y No frailejon



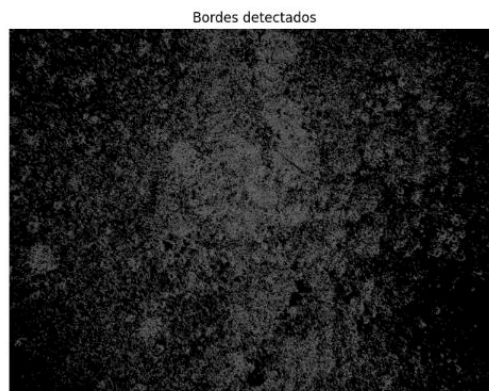
Se realizó un análisis de distribución de intensidades para determinar si hay una mayor presencia de ciertos colores en las imágenes de frailejones en comparación con las imágenes que no los contienen. En la imagen 2 de "Frailejón", la distribución de los colores parece más uniforme en comparación con las imágenes de "No Frailejón", donde hay una mayor concentración en ciertos rangos de intensidad. Los canales rojo, verde y azul presentan variaciones en su frecuencia relativa, lo que sugiere diferencias en la reflectancia y composición de color entre ambas categorías. Se evidencia que las imágenes del conjunto de datos Frailejón tienen una mayor presencia de tonos azul, en comparación con los datos de No frailejón. Las imágenes de No Frailejón muestran una mayor variabilidad en los tonos, especialmente en los canales rojo y azul, lo que podría reflejar la presencia de otros objetos o áreas del paisaje que no contienen vegetación densa.

Imagen 2. Detección de intensidades RGB de Frailejón y No frailejón



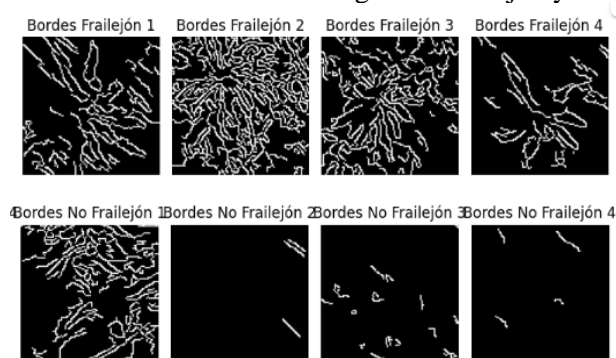
Por otra parte, se aplicó la técnica de detección de bordes a la imagen original completa para identificar áreas donde ocurren cambios abruptos en la intensidad de los píxeles, lo que permite resaltar los límites de los objetos presentes (ver imagen 3). La imagen muestra una gran cantidad de bordes, lo que indica una alta variabilidad en la textura del terreno y la vegetación. Sin embargo, los bordes detectados no presentan contornos claros de objetos, lo que sugiere una transición suave entre las áreas de la imagen.

Imagen 3. Bordes detectados en imagen original completa



En la imagen 4 de No Frailejón, los bordes detectados son en su mayoría finos y dispersos, lo que sugiere que las imágenes contienen estructuras sin contornos claramente definidos o que están más relacionadas con cambios graduales en la intensidad, mientras que las imágenes de Frailejón muestran bordes más densos y con formas más definidas, estas diferencias podrían ser útiles para la clasificación. Adicionalmente, se realizó un análisis de contornos para identificar las formas de los objetos presentes en la imagen (imagen 5).

Imagen 4. Bordes detectados en imagen de Frailejon y No frailejon



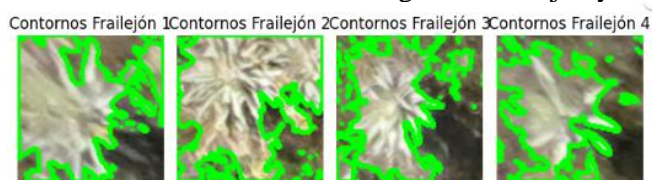
Los contornos resaltan principalmente las áreas con vegetación, lo que sugiere que estas zonas tienen una estructura más compleja y bien definida. En las áreas de menor vegetación, los contornos son menos prominentes (ver imagen 5).

Imagen 5. Contornos detectados en imagen original completa



En la Imagen 6 de Frailejón muestran contornos más numerosos y bien definidos alrededor de la estructura de las plantas. Los frailejones, por su morfología, tienen hojas alargadas y distintivas, lo que genera bordes más notables y complejos. Mientras que, en las de No Frailejón, aunque hay una detección de contornos, no son tan complejos o definidos como los de los frailejones. Los contornos parecen seguir patrones menos estructurados, lo que podría estar relacionado con elementos del paisaje, como rocas o terreno sin vegetación.

Imagen 6. Contornos detectados en imagen de Frailejón y No frailejón





## 5. Preparación de los datos

Para cumplir con los objetivos establecidos, se realizaron diversos procedimientos de preparación y preprocesamiento sobre las imágenes del conjunto de datos original. Estos procedimientos varían dependiendo del tipo de red que se va a entrenar. A continuación, se detallan las etapas más significativas:

**División de la imagen:** A partir de la imagen original, se procede a dividir el set en 250 imágenes de 70\*70 mediante la función 'Import\_Imagenes()' suministrada, estas imágenes se aplanaron para ser utilizadas en una red neuronal sencilla, estas imágenes contienen 14.700 características.

**Separación entre datos de entrenamiento y prueba:** En todos los casos, para el entrenamiento de los datos se dividió en dos grupos, conjunto de entrenamiento (CE) y conjunto de validación (CV). El conjunto de 250 imágenes se divide en una proporción de 70% para entrenamiento y 30% para prueba. De este modo, el conjunto de entrenamiento consta de 175 imágenes, mientras que el de prueba incluye 75 imágenes.

**Normalización de los valores RGB:** Las imágenes están en formato RGB, con valores que varían entre 0 y 255 para cada canal. Para normalizar estos valores, se dividen entre 255, lo que reescalará los datos al rango de 0 a 1, representados como valores decimales. Este proceso se aplica a todas las imágenes sin excepción.

**Reescalado de las imágenes:** Para los modelos convolucionales y los que emplean transferencia de aprendizaje se utilizaron las imágenes sin aplanar, fue necesario ajustar el tamaño de 70 x 70 píxeles, a 224 x 224 píxeles para su utilización modelos preentrenados como VGG16 o MobileNetV2. Por lo tanto, se realiza esta transformación mediante la función 'Resizing' de la librería TensorFlow en Python.

**Aumento de datos (Data Augmentation):** Se implementó un procedimiento de aumento de datos, mediante el cual se crean nuevas imágenes a partir de las originales. Este proceso incluye rotaciones en un rango de 20 grados, desplazamientos verticales y horizontales hasta un 20%, cambio el brillo de la imagen en un 20% e invierte la imagen con efecto espejo. El nuevo conjunto de datos resultante se utilizará para mejorar el rendimiento del modelo.

## 6. Modelos

A continuación, se describe brevemente cada una de las arquitecturas empleadas en los modelos solicitados para el desarrollo de los objetivos de negocio y minería de datos, en los cuales se hace uso de la función de pérdida binary\_crossentropy:

**Modelo 1- Red neuronal sencilla:** Para este modelo, la red neuronal cuenta con: una semilla fija (42), para asegurar que los resultados sean reproducibles. La red neuronal tiene un modelo secuencial, se define la red de manera lineal, capa por capa. Se tiene una capa oculta con 64 neuronas con función



de activación ReLU (Rectified Linear Unit), útil para evitar problemas de saturación en la propagación de gradientes. La capa de salida tiene 1 neurona con activación sigmoide, adecuada para clasificación binaria.

**Modelo 2 - Multi-capas:** Se establece una semilla (42) para asegurar que los resultados sean reproducibles. Se define y entrena un modelo de red neuronal con 3 capas ocultas, donde el número de neuronas en cada capa y otros parámetros (como la tasa de aprendizaje, el tamaño del lote y el número de épocas) son pasados como argumentos. Se compila el modelo con el optimizador Adam y la función de pérdida `binary_crossentropy`, se hizo una búsqueda de hiperparámetros dentro de los rangos: (neuronas 1 capa: 64, 128, 256), (neuronas 2 capa: 32, 64, 128), (neuronas 3 capa: 16, 32 y 64), tasa de aprendizaje 0.001, 0.1, 0.0001, tamaño de lote: 16, 32 y 64, épocas (50 y 100). Se selecciono aleatoriamente un subconjunto de combinaciones para acelerar el proceso de búsqueda (máximo 10 combinaciones). Teniendo en cuenta lo anterior la mejor configuración encontrada es: {'batch\_size': 64, 'epochs': 50, 'learning\_rate': 0.001, 'neuronas1': 64, 'neuronas2': 64, 'neuronas3': 16}, el proceso de búsqueda de hiper parámetros tomo un tiempo total de la búsqueda de: 195.51 segundos

**Modelo 3 - Convolutivo:** Se extrae dos conjuntos de imágenes: Frailejón y No Frailejón. Se define un modelo secuencial utilizando varias capas: tenemos 3 Capas convolucionales (Conv2D), seguidas de capas maxpooling2D, la diferencia es que las capas convolucionales tienen 32, 64 y 128 filtros respectivamente con kernel de 3x3 con activación ReLU y Dropout de 0,5 y early stopping para prevenir el sobreajuste, posteriormente una capa de aplanamiento (Flatten) para convertir las características en un vector, una capa de salida con activación sigmoide para la clasificación binaria y se compila el modelo con el optimizador Adam.

**Modelo 4 - Convolutivo + VGG16:** Se carga la arquitectura VGG16 preentrenada en el conjunto de datos ImageNet. Se establece `include_top=False` para excluir las capas completamente conectadas, ya que se agregarán nuevas capas para la clasificación específica del problema. Las capas convolucionales de VGG16 se congelan (`base_model.trainable = False`), evitando que se reentrenen durante el ajuste fino, lo cual es útil para aprovechar las características aprendidas sin sobre ajustar los parámetros. Se construye un modelo secuencial que incluye: La base VGG16, una capa de aplanamiento (Flatten) para convertir las salidas en un vector, una capa densa (Dense) con 128 neuronas y activación ReLU, una capa de Dropout 0.5 para prevenir el sobreajuste y una capa de salida con activación sigmoide para la clasificación binaria.

**Modelo 5 - Convolutivo con transferencia VGG-16 y ajuste de hiperparámetros:** tomando como referencia la red neuronal VGG16 se procede a hacer un ajuste de hiper parámetros con el uso de la librería Optuna, se prueban los parámetros: neuronas en la capa densa 64 y 512, dropout de 0,3 a 0,7 en pasos de 0,1, optimizador “Adam”, “sgd”, “rmsprop” y learning rate entre  $1e-5$  y  $1e-2$ . Teniendo en cuenta lo anterior, los mejores hiperparámetros encontrados son: {'neuronas': 384, 'dropout': 0.4, 'optimizer': 'adam', 'learning\_rate': 0.00418397}.

**Modelo 6- MobileNetV2** esta es otra red neuronal, pero pre-entrenada distinta a la VGG-16. Se utiliza MobileNetV2 como la base del modelo, cargando pesos preentrenados en el conjunto de datos ImageNet. Se establece `include_top=False`, lo que significa que no se incluirán las capas finales de clasificación original del modelo, permitiendo personalizar la parte superior. Se incluye una capa GlobalAveragePooling2D se agrega esta capa para reducir la dimensionalidad del output del modelo base, realizando un promedio global sobre las características extraídas. Esto ayuda a evitar el sobreajuste y reduce la cantidad de parámetros en el modelo. En las capas densas, primera capa densa: Contiene 512 neuronas y utiliza la función de activación ReLU. En la parte de Dropout se aplica una

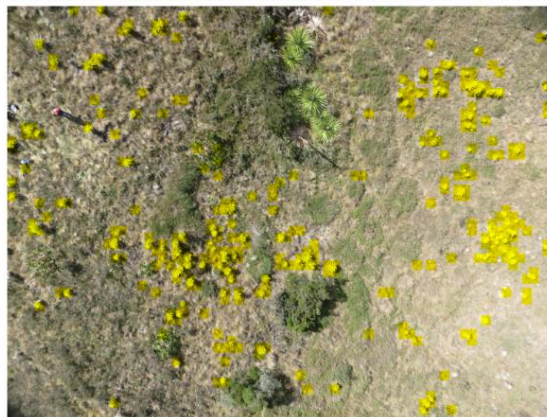
capa de dropout con una tasa de 0.4 para prevenir el sobreajuste durante el entrenamiento. En la segunda capa densa: contiene 256 neuronas y también utiliza la función de activación ReLU. En el dropout se aplica otra capa de dropout, esta vez con una tasa de 0.3. Por último, la capa de salida contiene 1 neurona y utiliza una función de activación sigmoide. El modelo usa el optimizador Adam con una tasa de aprendizaje de  $1e-4$ .

MobilnetV2 es una arquitectura de red neuronal diseñada para la clasificación de imágenes, especialmente enfocada en dispositivos móviles, aplicaciones con recursos limitados e imágenes pequeñas teniendo en cuenta que manejamos imágenes pequeñas (70x70). Fue desarrollada por Google y se caracteriza por su eficiencia en términos de tamaño y velocidad, manteniendo un buen rendimiento en tareas de reconocimiento de imágenes, es una opción apropiada en la detección de frailejones pues ha aprendido a diferenciar características asociadas a texturas, bordes, formas, etc, que pueden ser reutilizadas basados en la naturaleza del proyecto. De esta forma, los patrones visuales en la tarea base, asociada a la clasificación de imágenes en ImageNet (contiene más de 14 millones de imágenes etiquetadas, organizadas en más de 20,000 categorías) y la tarea objetivo, que comprende la detección de frailejones, se encuentran fuertemente relacionadas; así, el mapa de características aprendidas en la red neuronal fundamenta un modelo de aprendizaje valioso para la necesidad planteada.

#### **Evaluación cualitativa del mejor modelo:**

En la imagen 7, se observa la predicción del número de frailejones sobre la imagen ‘IMG\_3451.JPG’, para el desarrollo e inspección del modelo y su aplicación para reconocimiento de densidad poblacional de frailejones en el páramo de Chingaza y Cruz Verde. Los recuadros amarillos muestran la predicción de frailejón sobre la imagen general. Se observa que el modelo 5 Convolutacional con transferencia VGG-16 y ajuste de hiperparámetros es exitoso para determinar las zonas donde existen frailejones e identificarlos. En comparación con los otros modelos evaluados es el mejor en la identificación de frailejones, observando en la imagen 7 a todos los frailejones con puntos amarillos, lo cual facilita el cálculo automático de la densidad poblacional en el páramo por medio de reconocimiento de imágenes.

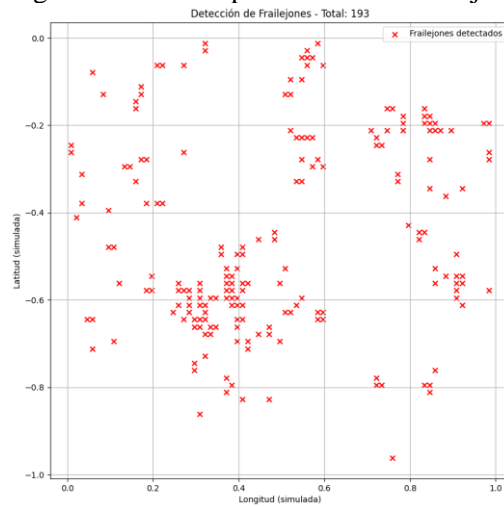
Imagen 7. Identificación cualitativa del mejor modelo resultante



#### **Cálculo de densidad poblacional:**

El modelo seleccionado encontró satisfactoriamente 193 frailejones sobre la imagen de dimensiones 3000 x 4000 pixeles (ver imagen 8).

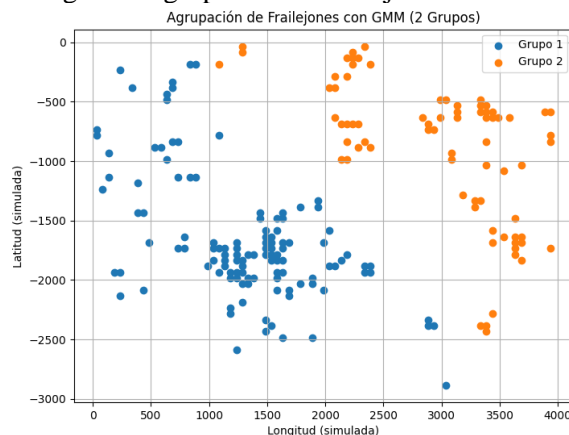
Imagen 8. Densidad poblacional de frailejones



Se aplica Gaussian Mixture Model (GMM) para agrupar las detecciones de frailejones en 2 clúster y por otra parte, se calcula la densidad poblacional de cada grupo. Se define `num_clusters = 2`, lo que indica que se agruparán las detecciones en 2 grupos. Se aplica GMM a las coordenadas de los frailejones detectados. Se obtiene la etiqueta de cluster para cada detección. Finalmente, se realiza el cálculo del número total de cuadros en los que se dividió la imagen, se cuenta cuántos cuadros contienen frailejones en cada grupo, se calcula la densidad como la proporción de cuadros con frailejones sobre el total de cuadros y el resultado se visualiza en la imagen 9 donde cada cluster se representa con un color diferente. A continuación, se muestran los resultados de la densidad poblacional por cada grupo:

- Densidad en Grupo 1:  $118 / 4661 = 0.025316$  cuadros con frailejones/cuadro total
- Densidad en Grupo 2:  $75 / 4661 = 0.016091$  cuadros con frailejones/cuadro total
- Densidad Total:  $193 / 4661 = 0.041407$  cuadros con frailejones/cuadro total

Imagen 9. Agrupación de frailejones con GNN



Como conclusión del ejercicio, se demuestra que el desarrollo de modelos de reconocimiento de imágenes es útil para detectar zonas de baja densidad poblacional, y de esta forma, que éstas sean objeto para estrategias de reforestación. Los metadatos de cada imagen contienen sus coordenadas



espaciales por lo que las autoridades competentes ya pueden contar con herramientas de detección preventiva y automatizada. Lo anterior da respuesta al objetivo de negocio planteado originalmente.

## 7. Evaluación

A continuación, se muestran los resultados de los modelos teniendo en cuenta la métrica indicada (AUC):

Nombre del modelo	AUC en entrenamiento	AUC en validación	AUC objetivo
Modelo 1 - Red neuronal sencilla	0.98	0.88	0.88
Modelo 2 - Multi-capa	0.81	0.89	0.95
Modelo 3 - Convolutacional	0.96	0.96	0.95
Modelo 4 - Convolutacional + VGG16	0.99	0.99	0.99
Modelo 5 - Convolutacional con transferencia VGG-16 y ajuste de hiperparámetros	1.00	1.00	No aplica
Modelo 6 - MobileNetV2	0.98	0.82	No aplica

El Modelo 1, que es una red neuronal sencilla, tiene un AUC de 0.88, lo que indica un rendimiento razonable pero no óptimo. El Modelo 2, una red multi-capa, mejora ligeramente con un AUC de 0.89, lo que sugiere que añadir más capas puede ayudar a capturar mejor la complejidad de los datos. Los Modelos 3 y 4, que son convolucionales, muestran un AUC significativamente más alto, con 0.96 y 0.99, respectivamente. Esto indica que las redes convolucionales son mucho más efectivas para este tipo de tarea, probablemente debido a su capacidad para extraer características espaciales de las imágenes.

El Modelo 5, que utiliza la transferencia de aprendizaje con VGG16 y ajuste de hiperparámetros, alcanza un AUC perfecto de 1.00. Esto sugiere que la combinación de un modelo preentrenado y la optimización de hiperparámetros puede llevar a un rendimiento sobresaliente. Este modelo demuestra la efectividad del transfer learning, especialmente en tareas de clasificación de imágenes donde hay un conjunto limitado de datos. El Modelo 6, que utiliza MobileNetV2, presenta un AUC de 0.82, que es el más bajo entre los modelos convolucionales. Esto podría deberse a diferencias en la arquitectura o a cómo se entrenó el modelo en comparación con los otros.

Para finalizar los mejores resultados, se presentan en las redes convolucionales (Modelos 3 y 4) son claramente más efectivas que las redes neuronales sencillas o multi-capa para tareas de clasificación de imágenes. La transferencia de aprendizaje con un modelo preentrenado como VGG16 (Modelo 5) puede llevar a mejoras significativas en el rendimiento, incluso alcanzando un AUC perfecto. Aunque MobileNetV2 es un modelo más sofisticado, no siempre garantiza un mejor rendimiento en todas las tareas. Puede requerir un ajuste de hiperparámetros más fino o una mejor adaptación a los datos específicos del problema. Los resultados sugieren que la optimización de hiperparámetros es crucial para alcanzar un rendimiento óptimo en modelos más complejos.

## Referencias

Instituto Humboldt. (6 de enero de 2022). *Nueva especie de frailejón en Colombia, 92 especies y contando...* Instituto Humboldt. <https://www.humboldt.org.co/noticias/nueva-especie-de-frailejon-en-colombia-92-especies-y-contando>