



# Dermatología con IA

## Introducción

En este proyecto, se entrenó una inteligencia artificial con el objetivo de identificar enfermedades de la piel mediante el análisis de ciertos parámetros. Se presentará la metodología empleada, los mejores resultados obtenidos, las conclusiones alcanzadas y las posibles mejoras a considerar en el futuro.

## Descripción del Dataset

El conjunto de datos que utilizamos para el entrenamiento cuenta con 34 características (parámetros a usar) y 366 instancias (cantidad de datos disponibles). Además, tenemos una característica llamada 'clase', que indica la enfermedad correspondiente según los otros parámetros. Las características están divididas entre atributos clínicos y atributos histopatológicos.

A partir de los datos, podemos inferir lo siguiente: no existen datos de tipo "object", por lo que no será necesario mapearlos. La edad está fuera del rango del resto de los parámetros, por lo que debemos normalizarla. Contamos con 34 entradas (inputs) y una clasificación, y hay 6 salidas (outputs) correspondientes a las 6 clases. En la página del dataset, se observó que una clase contiene valores faltantes (missing values), por lo que deben ser corregidos. Además, las clases no están bien balanceadas, por lo que debemos crear datos sintéticos para equilibrarlas.

## Resultados

Modelo con mejor resultado:

```
modelo1 = tf.keras.models.Sequential([
    tf.keras.layers.InputLayer(input_shape=(34)),
    tf.keras.layers.Dense(units=32, activation='relu'),
    tf.keras.layers.Dense(units=32, activation='relu'),
    tf.keras.layers.Dense(units=32, activation='relu'),
    tf.keras.layers.Dense(units=6, activation='softmax')
])

# Compilando
modelo1.compile(loss=tf.keras.losses.SparseCategoricalCrossentropy(), optimizer='adam', metrics=['accuracy'])
```

Accuracy

```
[ ] modelo1.evaluate(x=test_set, y=test_target, batch_size=128)
```

```
2/2 [=====] - 0s 7ms/step - loss: 0.0363 - accuracy: 1.0000
[0.03630714491009712, 1.0]
```

Confusion Matrix

| True label | Predicted label |    |    |    |    |    |
|------------|-----------------|----|----|----|----|----|
|            | 0               | 1  | 2  | 3  | 4  | 5  |
| 0          | 31              | 0  | 0  | 0  | 0  | 0  |
| 1          | 0               | 23 | 0  | 0  | 0  | 0  |
| 2          | 0               | 0  | 18 | 0  | 0  | 0  |
| 3          | 0               | 0  | 0  | 25 | 0  | 0  |
| 4          | 0               | 0  | 0  | 0  | 20 | 0  |
| 5          | 0               | 0  | 0  | 0  | 0  | 17 |

## Metodología

Lo primero a realizar en el preprocesamiento es la eliminación de los valores faltantes. Dado que el dataset indica que existen estos valores, optamos por eliminarlos ya que no suelen aportar nada útil.

Se realiza una normalización, que consiste en convertir valores que tienden a tener una gran variación en valores entre 0 y 1. Esto facilita el análisis de los datos.

Luego, revisamos la cantidad de instancias por cada clase, como se muestra en la figura 1. Es recomendable realizar un balanceo de clases, lo cual implementamos mediante el oversampling. Esto hace que la cantidad de instancias de todas las clases sea igual a la de la clase de psoriasis, como se puede visualizar en la figura 2.

|   |                          |     |
|---|--------------------------|-----|
| 1 | psoriasis                | 112 |
| 2 | seboric dermatitis       | 61  |
| 3 | lichen planus            | 72  |
| 4 | pityriasis rosea         | 49  |
| 5 | ronic dermatitis         | 52  |
| 6 | pityriasis rubra pilaris | 20  |

Figura 1

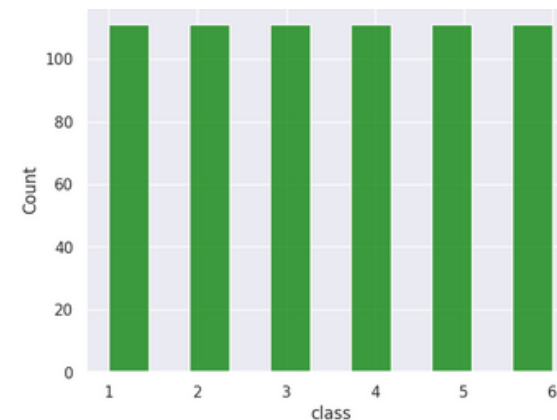


Figura 2

## Conclusiones

En conclusión, este proyecto nos mostró diferentes formas de crear un modelo para identificar enfermedades de la piel. La calidad y precisión de los datos son cruciales, aunque no siempre son perfectos. También aprendimos sobre la evaluación médica en dermatología y el impacto revolucionario de la inteligencia artificial en la medicina.

## Mejoras a futuro

Trabajar con más datos y de diversas fuentes para mejorar la representatividad y robustez del modelo. A demás, crear nuevas características a partir de las existentes con Feature Engineering. Ajustar hiperparámetros exhaustivamente y Ensemble Learning para combinar varios modelos para mayor precisión y robustez.



## Integrantes:

**Cristopher Azurdia**  
-21000232-  
**Ramón Rivas**  
-20000665-  
**José Barrios**  
-21000478-