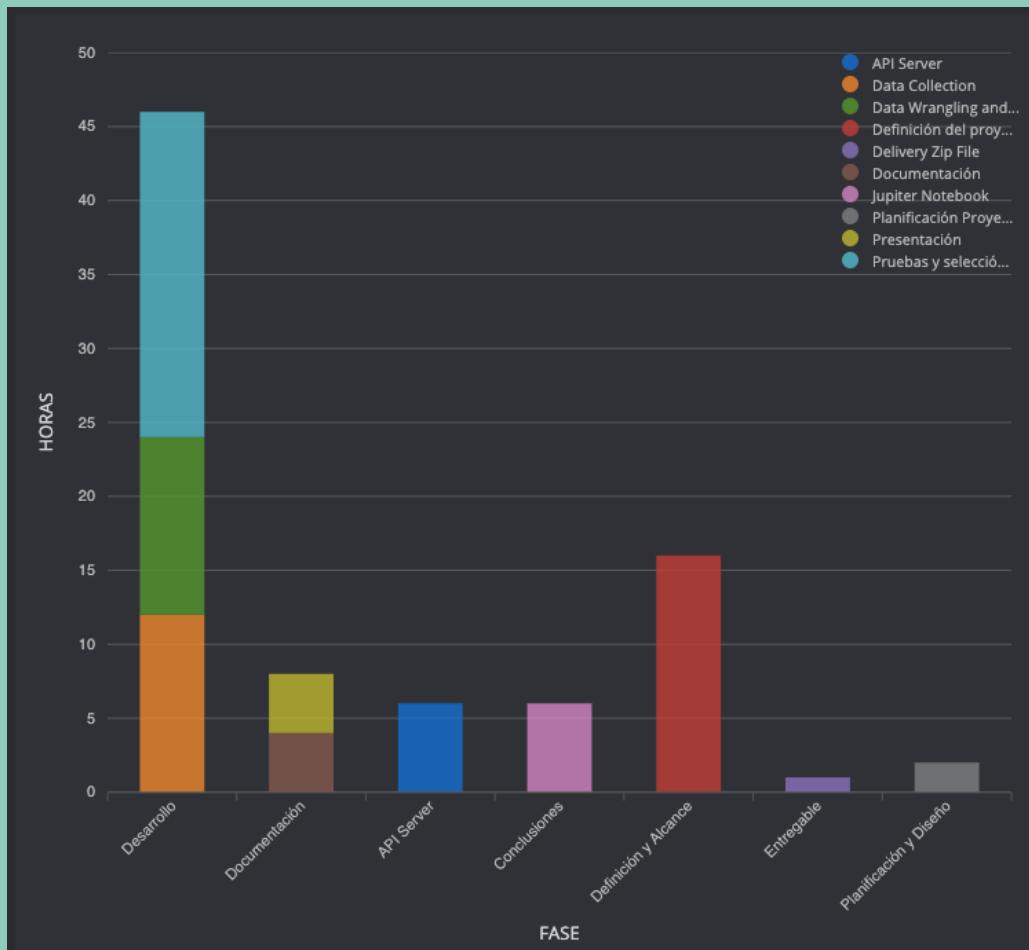


THE BRIDGE - DATA SCIENCE BOOTCAMP

LESIONES CUTÁNEAS ANALÍTICA AVANZADA- DEEP LEARNING

MÓNICA VILLASUSO LÓPEZ - OCTUBRE 2020

85 HORAS



Contexto

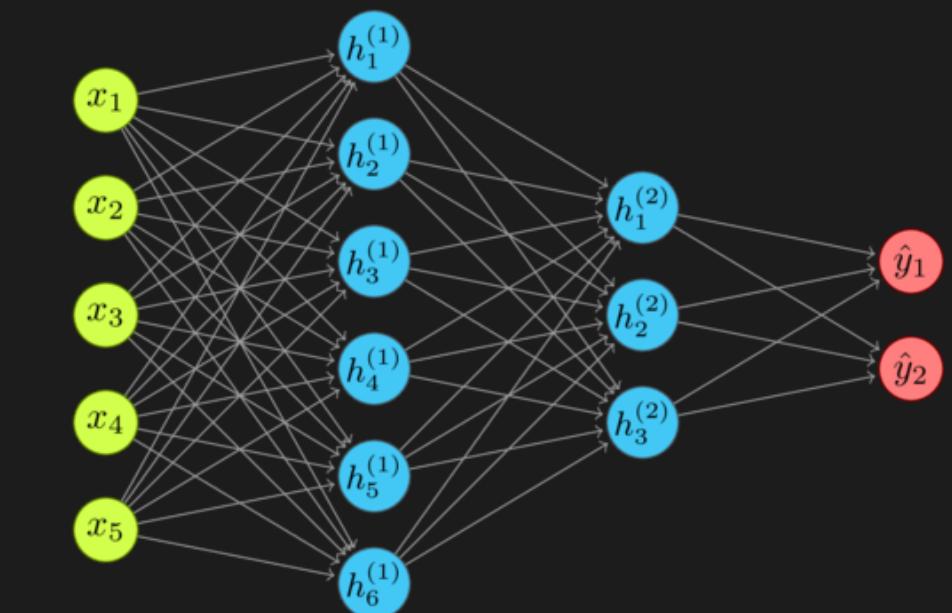
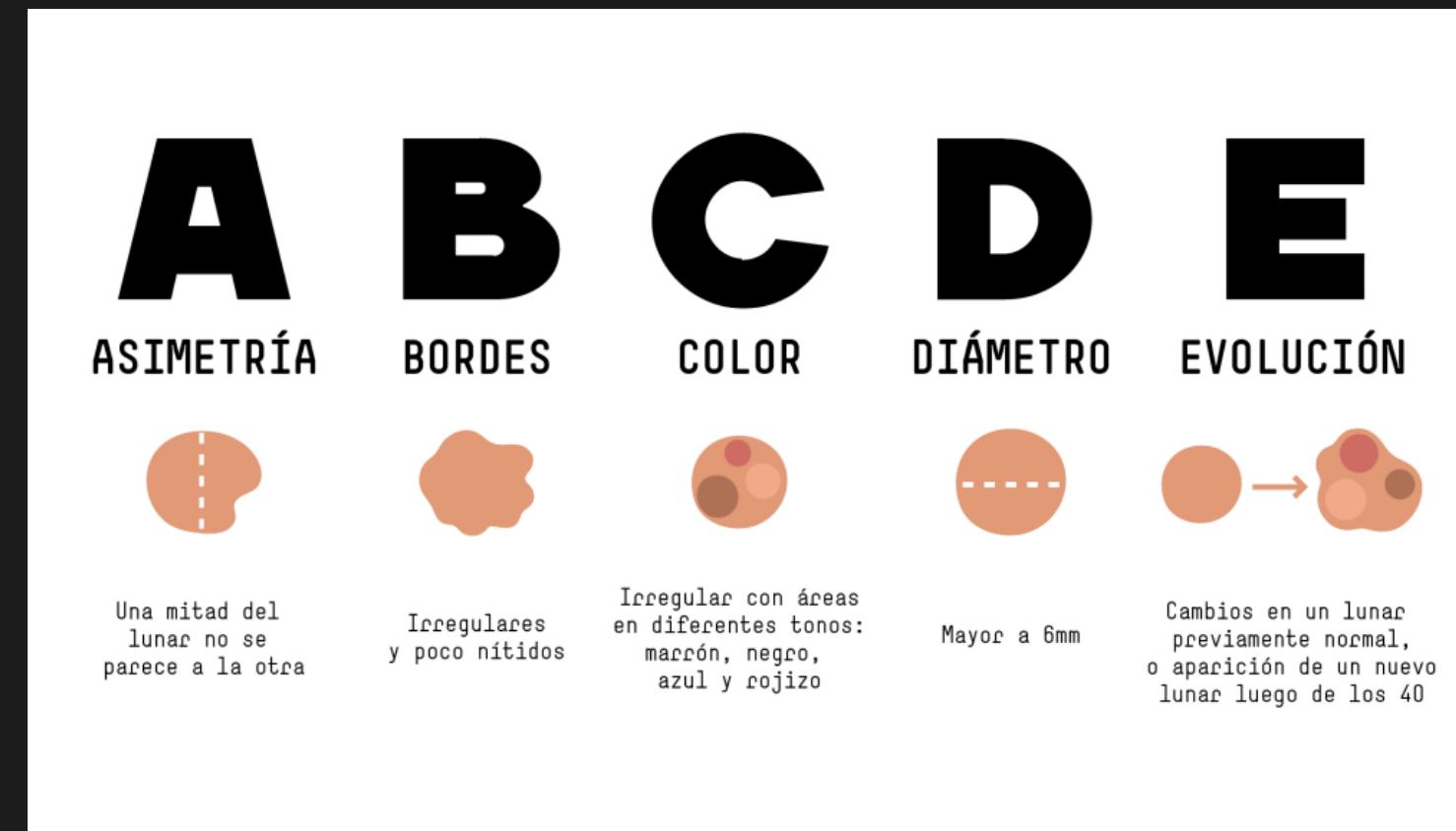
Recolección de datos

Tratamiento de datos

Entrenamiento del modelo

Resultados y Demo (API)

ANTECEDENTES - CONTEXTO



CLASIFICACIÓN DE IMÁGENES

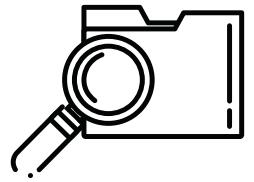
- SALUD -

DEEP LEARNING



- En una década los casos globales de melanoma han aumentado un 50%, situándose en 287.000 casos anuales en todo el mundo.
- Para 2025 se espera un crecimiento del 18%.
- En España cada año se diagnostican 5.000 casos de melanoma y 71.000 casos de cáncer de piel no melanoma

RECOLECCIÓN DE DATOS



Recolección

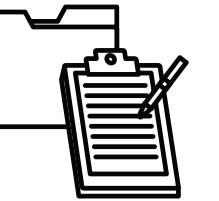
303 imágenes



Organización

142 - Malignas

161 - Benignas

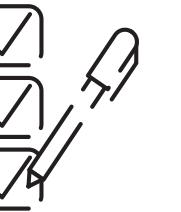


Pruebas de concepto

Color / B&W

Dimensiones (varias)

Normalización o no

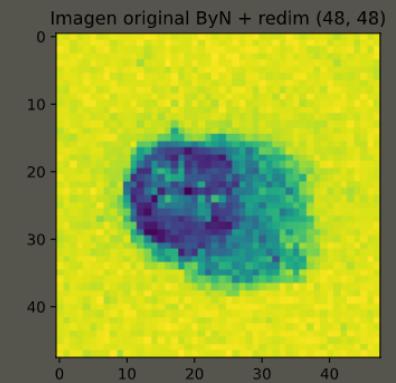
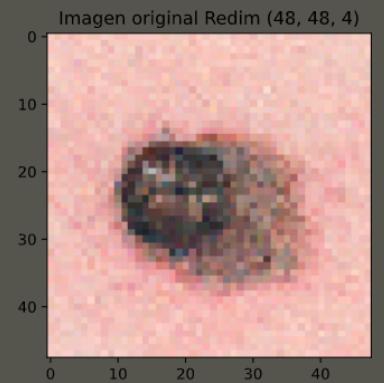
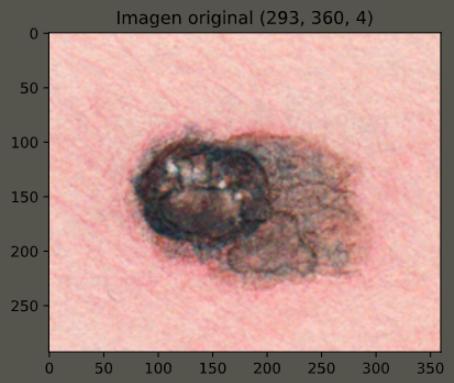
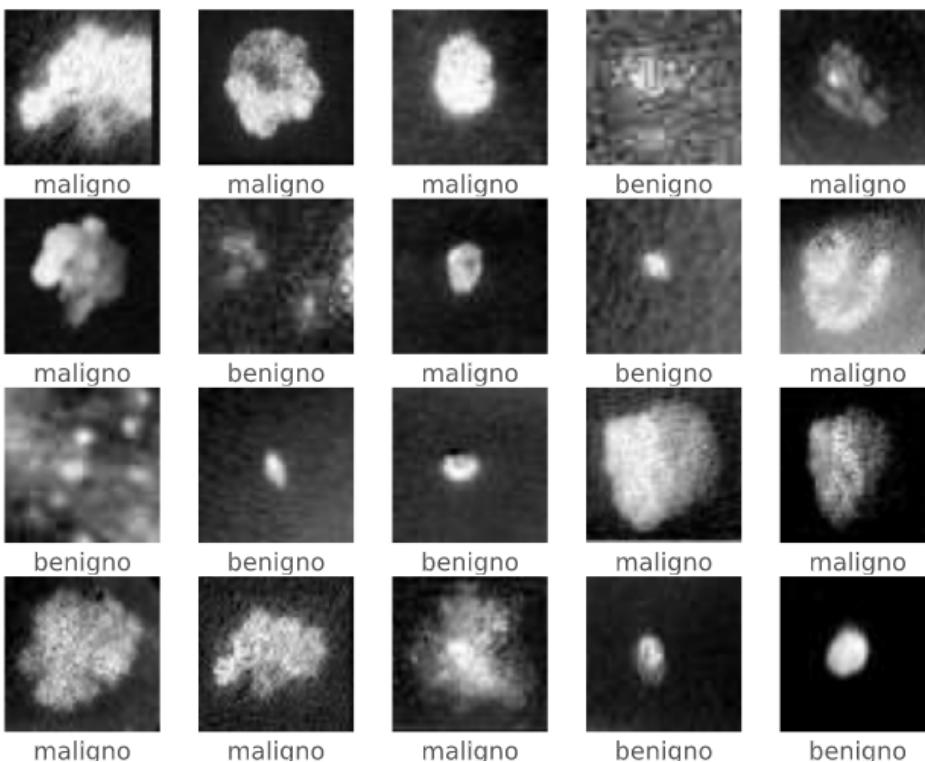


Imágenes

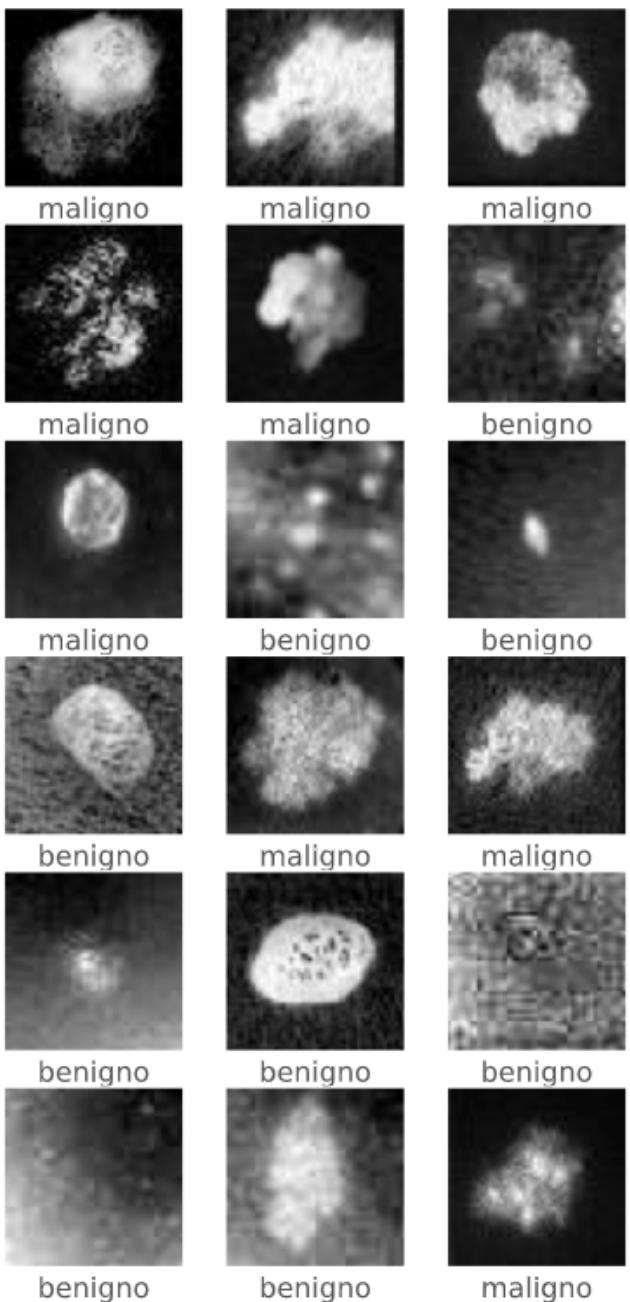
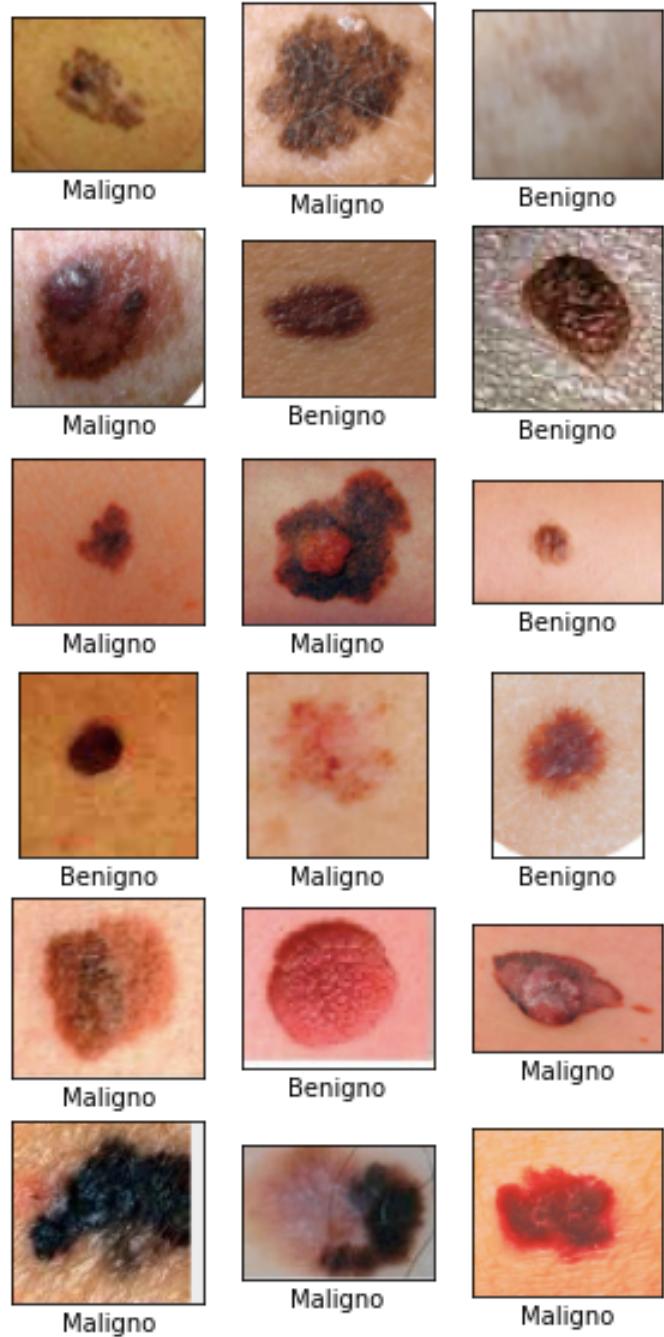
B&W

48x48

Normalizadas

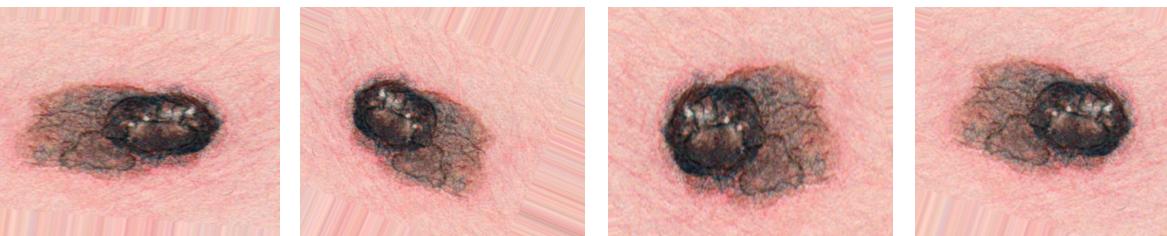


TRATAMIENTO DE DATOS



IMAGEDATAGENERATOR - KERAS
303 --> 10.000

(horizontal_flip=True,
rotation_range=20,
width_shift_range=0.1,
height_shift_range=0.1,
zoom_range=0.2,
shear_range=0.1,
fill_mode='nearest')



PARTICIÓN DE CONJUNTOS

TRAIN - 85%
TEST - 15%

CROSS VALIDATION
(VALID= 20% TRAIN)

ENTRENAMIENTO Y ENTONACIÓN DEL MODELO

Convolutional Neural Network (CNN)
EPOCHS = 100 - BATCH SIZE = 48

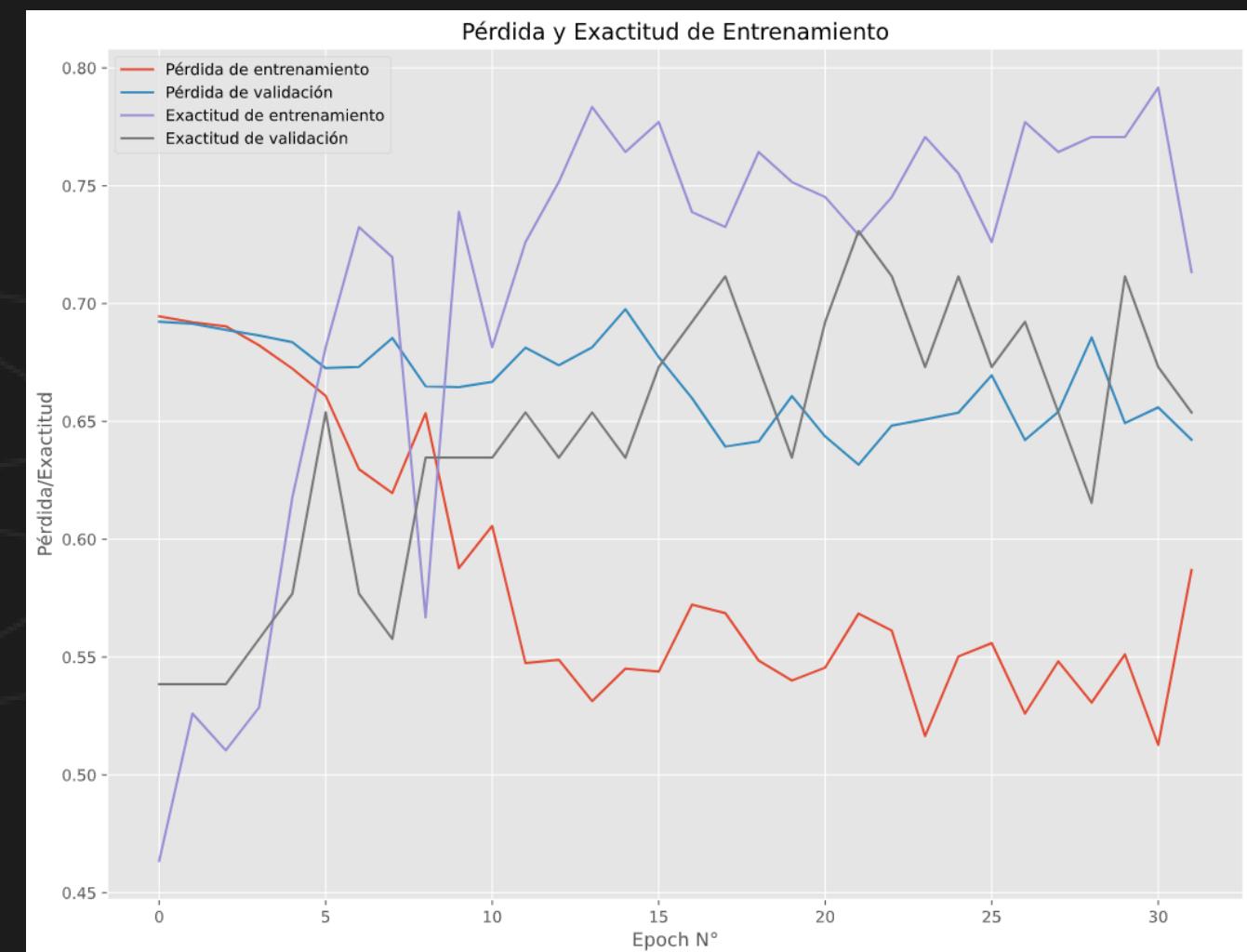
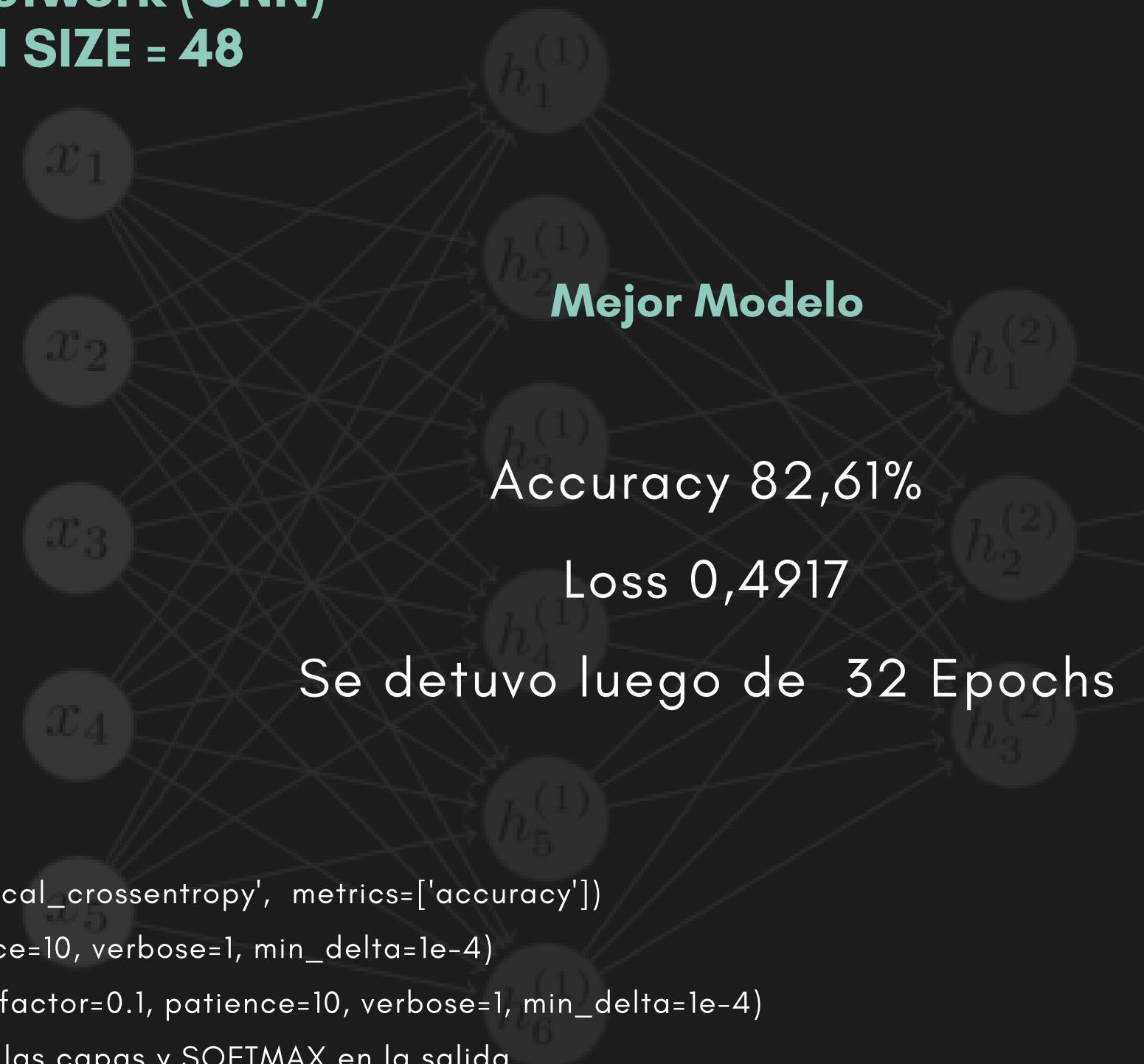
L 1 capa convolutiva con 8 neuronas
 - 1 capa convolutiva con 16 neuronas
 - 1 MaxPool
 - 1 dropout 0.25
 - 1 capa convolutiva con 32 neuronas
 - 1 MaxPool
 - 1 dropout 0.25
 - 1 Flatten
 - 1 dense con 32 neuronas
 - 1 dense con 64 neuronas
 - 1 dense con 2 (Softmax --> salida)

optimizer='Adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])

EarlyStopping(monitor='val_loss', patience=10, verbose=1, min_delta=1e-4)

ReduceLROnPlateau (monitor='val_loss', factor=0.1, patience=10, verbose=1, min_delta=1e-4)

Función de activación = RELU en todas las capas y SOFTMAX en la salida



Otras pruebas

Activación: RELU- variaciones, TanH

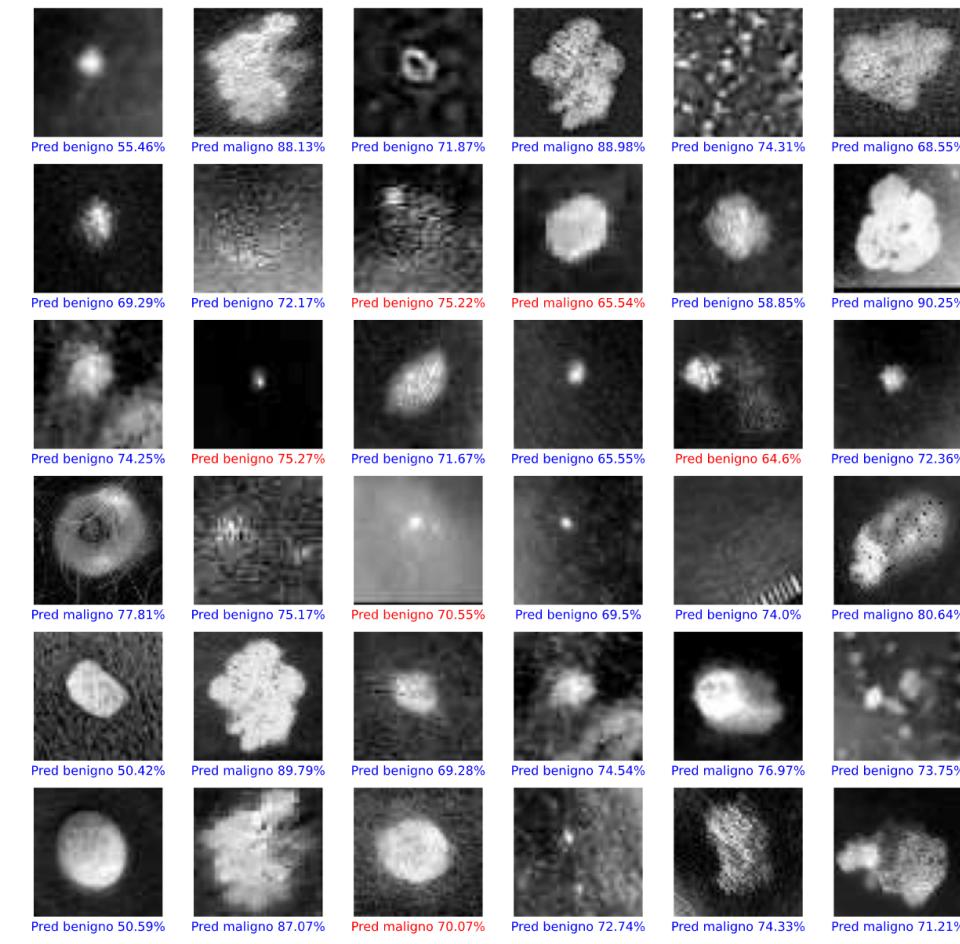
Optimizador: RMSprop, SGD, Adadelta, Adagrad, Adamax

RESULTADOS

Accuracy 82,61%

Loss 0,4917

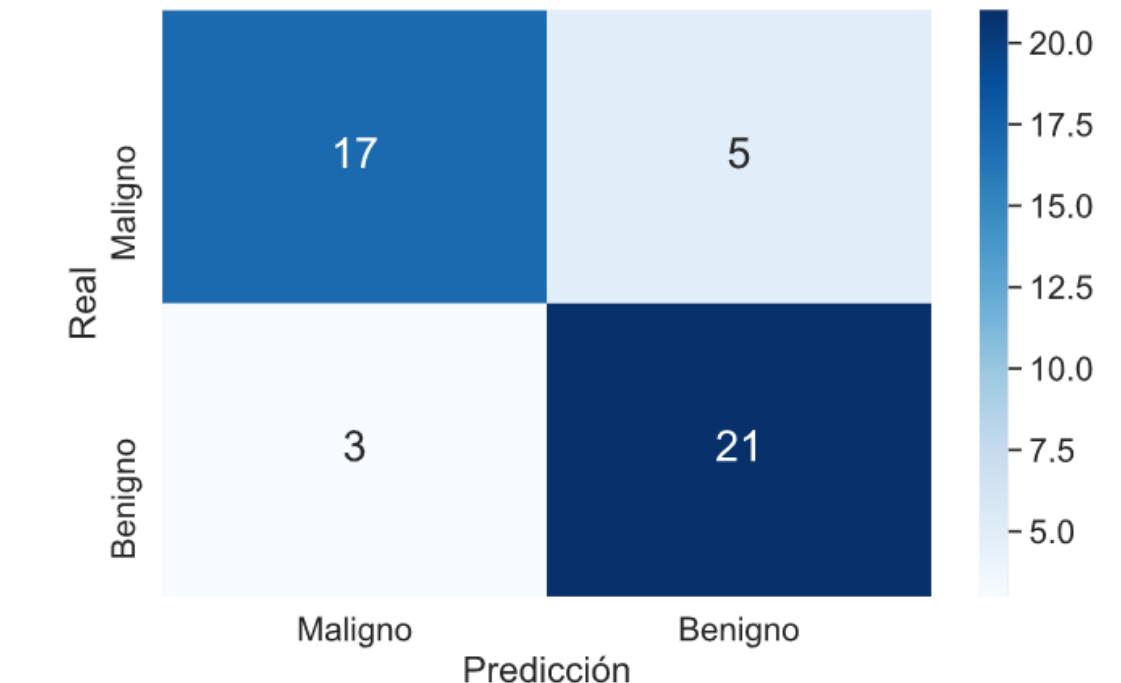
Classification Report				
	precision	recall	f1-score	support
Maligno	0.85	0.77	0.81	22
Benigno	0.81	0.88	0.84	24
accuracy			0.83	46
macro avg	0.83	0.82	0.82	46
weighted avg	0.83	0.83	0.83	46



Reporte de Clasificación

Resultados conjunto Test

MATRIZ DE CONFUSIÓN



Matriz de confusión

PREDICCIÓN DEL MODELO

CASO REAL



<http://localhost:6060/get/pred?mod=vista>



API (FLASK Server)

2 end points (Visual / archivo Json)

Resultado del modelo [json]
{"Sospechoso":{"0":0.6626951098}, "Benigno":{"0":0.3373049498}}

<http://localhost:6060/get/pred?mod=json>

GROUND TRUTH

CASO REAL



“

A LOS MAYORES LES GUSTAN LAS CIFRAS.
CUANDO SE LES HABLA DE UN NUEVO AMIGO, JAMÁS
PREGUNTAN SOBRE LO ESENCIAL DEL MISMO.

NUNCA SE LES OCURRE PREGUNTAR:

"¿QUÉ TONO TIENE SU VOZ?

¿QUÉ JUEGOS PREFIERE?

¿LE GUSTA COLECCIONAR MARIPOSAS?".

PERO EN CAMBIO PREGUNTAN:

"¿QUÉ EDAD TIENE?

¿CUÁNTOS HERMANOS?

¿CUÁNTO PESA?

¿CUÁNTO GANA SU PADRE?"

SOLAMENTE CON ESTOS DETALLES CREEN CONOCERLE.

El Principito - Antoine de Saint-Exupéry

”

GRACIAS

THE BRIDGE
DATA SCIENCE BOOTCAMP
JUN - OCT | 2020

