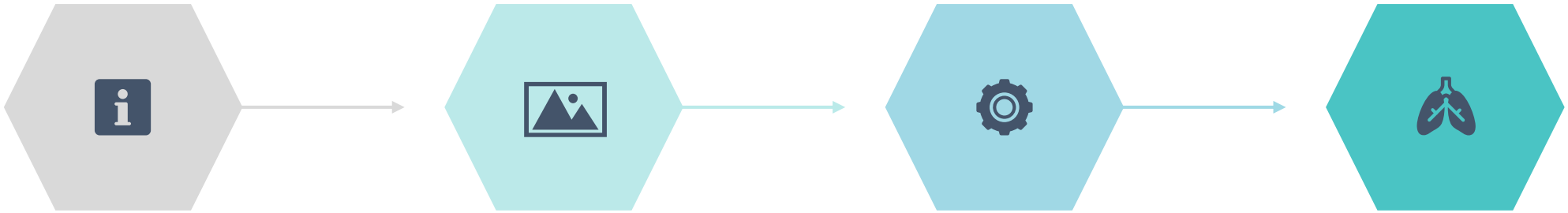
Diagnóstico de neumonía

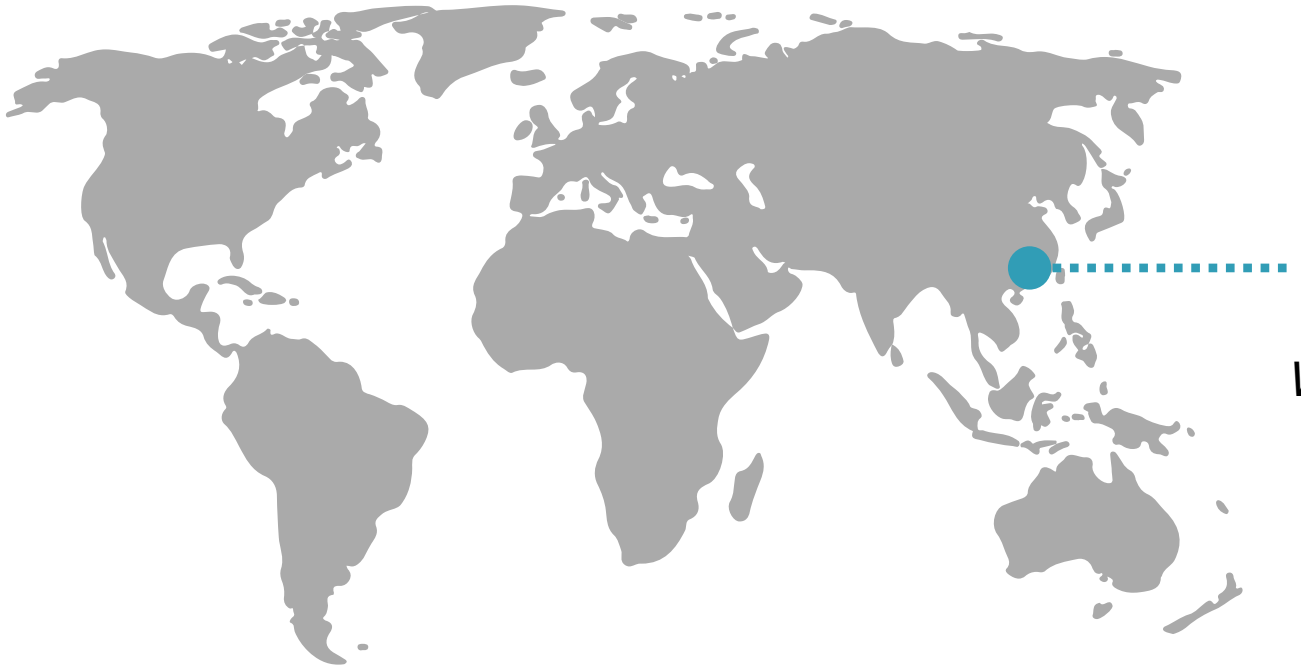
Clasificador binario

Contenidos



Datos **Imágenes** **Modelos** **Resultados**

Chest X-Ray Images (Pneumonia)

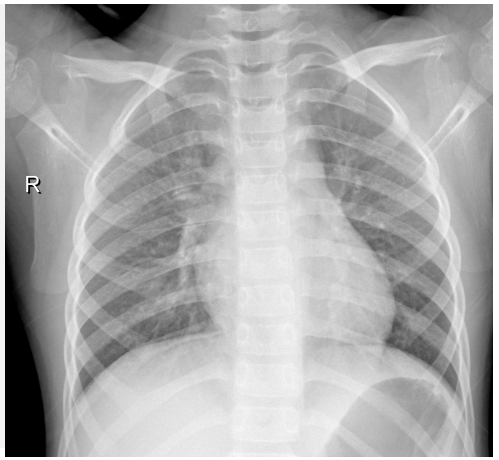


Guangzhou, China

Women and Children's Medical Center

Imágenes

5856 archivos

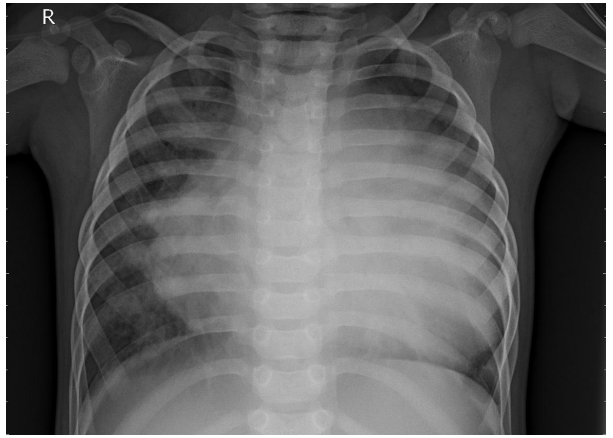


Normal

1341 train

234 test

8 validation

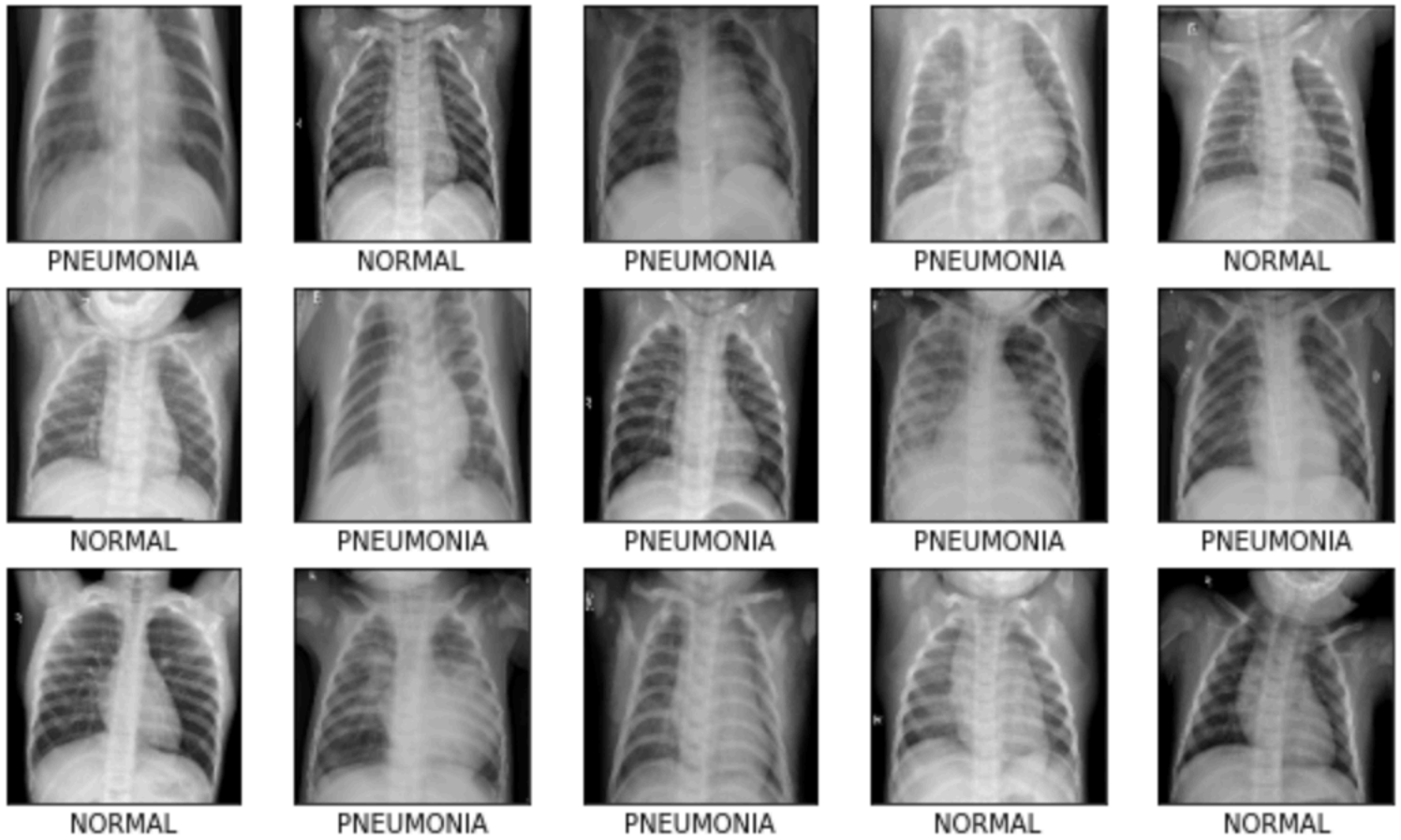


Neumonía

3875 train

390 test

8 validation



Shape

X_train: (5216, 128, 128, 3)

y_train: (5216,)

X_test: (624, 128, 128, 3)

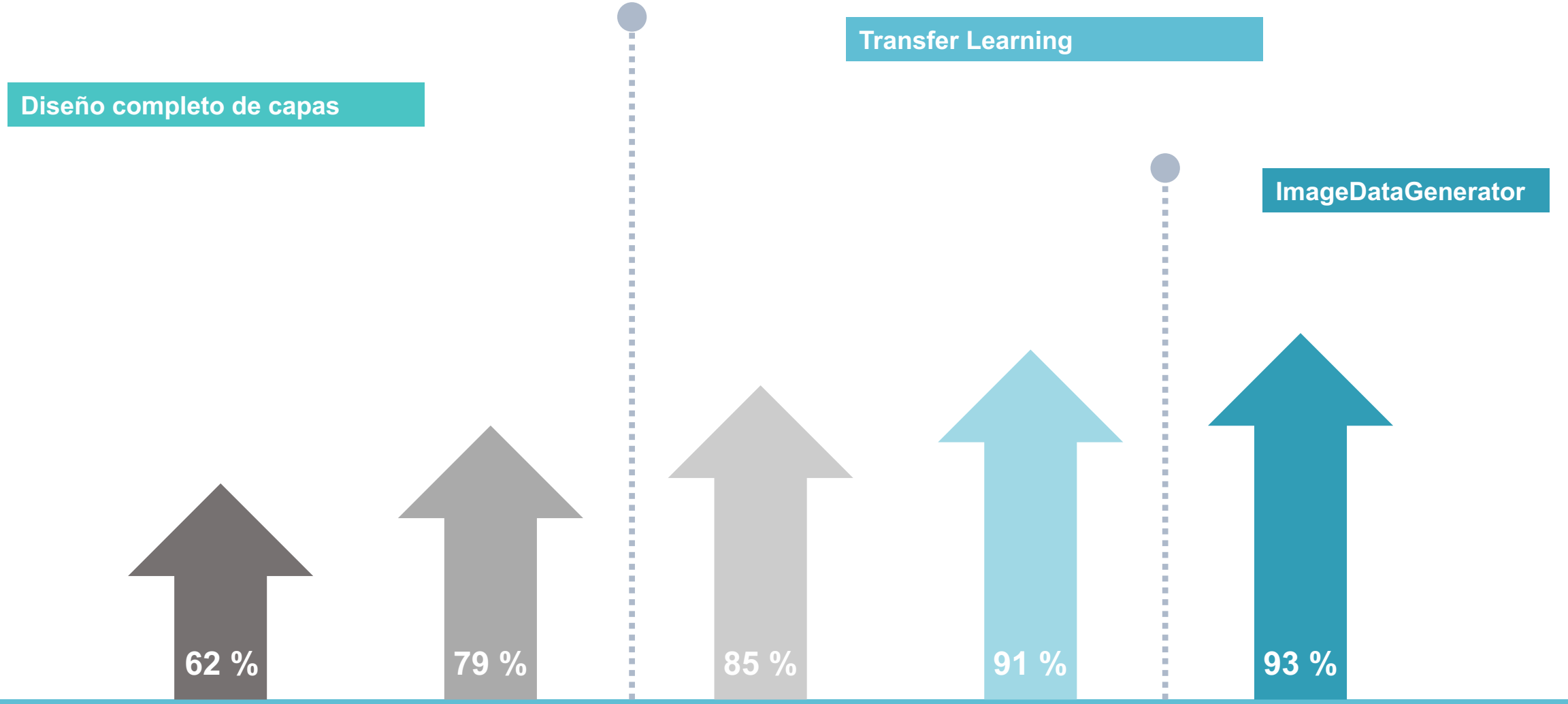
y_test: (624,)

Normalización

X_train = X_train / 255.0

X_test = X_test / 255.0

X_val = X_val / 255.0



Modelos

128x128 | tensorflow | keras | deep learning

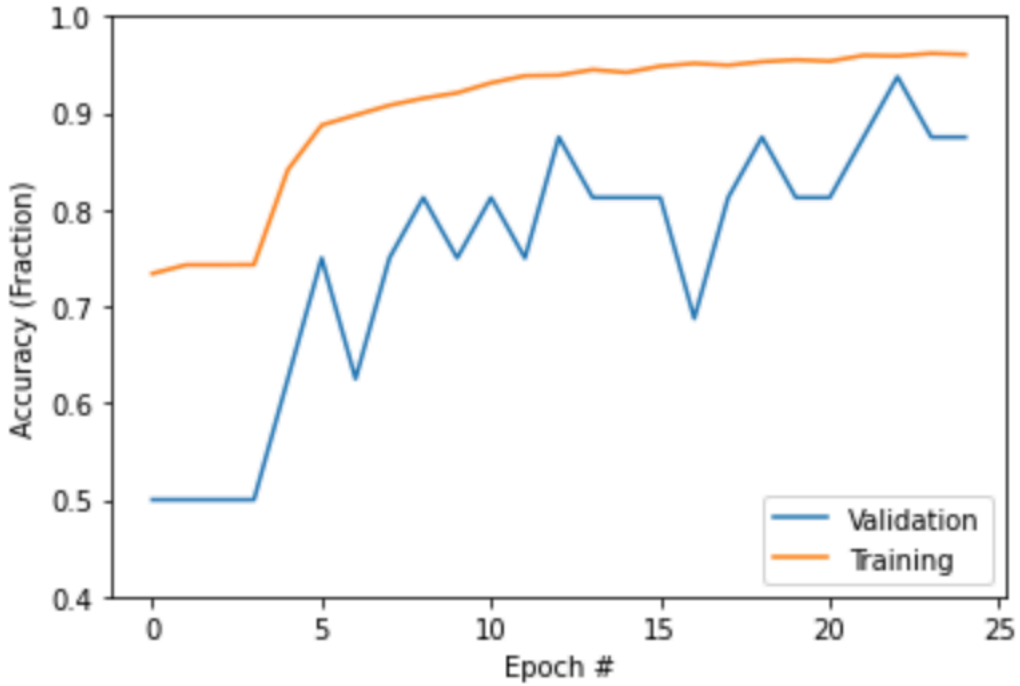
VGG16

93 %

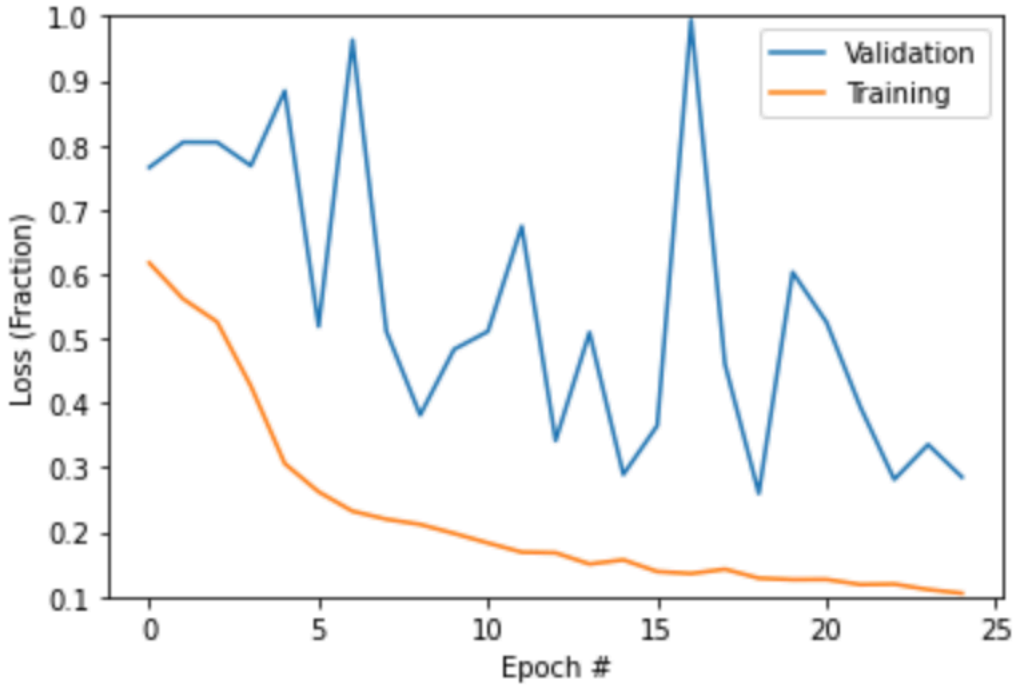
```
# Arquitectura

expert_conv = VGG16(weights = 'imagenet', include_top = False,input_shape=(128,128,3))

for layer in expert_conv.layers:
      trainable = True
      layer.trainable = trainable

tweaked_model = Sequential()
tweaked_model.add(Reshape((128,128,3)))
tweaked_model.add(expert_conv)
tweaked_model.add(GlobalAveragePooling2D())

tweaked_model.add(Dense(128, activation = 'relu'))
tweaked_model.add(Dropout(0.3))
tweaked_model.add(Dense(4096, activation = 'relu'))
tweaked_model.add(Dense(1, activation = "sigmoid"))

opt = keras.optimizers.SGD(lr=1e-4, momentum=0.8)

tweaked_model.compile(loss = "binary_crossentropy", optimizer = opt, metrics=['accuracy'])

history = tweaked_model.fit(datagen.flow(X_train,y_train, batch_size = 32) ,epochs = 25,
                            validation_data = validation_datagen.flow(X_val, y_val))
```

```
datagen = ImageDataGenerator(
        featurewise_center=False,
        samplewise_center=False,
        featurewise_std_normalization=False,
        samplewise_std_normalization=False,
        zca_whitening=False,
        rotation_range = 30,
        zoom_range = 0.2,
        width_shift_range=0.1,
        height_shift_range=0.1,
        horizontal_flip = True,
        vertical_flip=False)

validation_datagen = ImageDataGenerator()

datagen.fit(X_train)
```

Accuracy



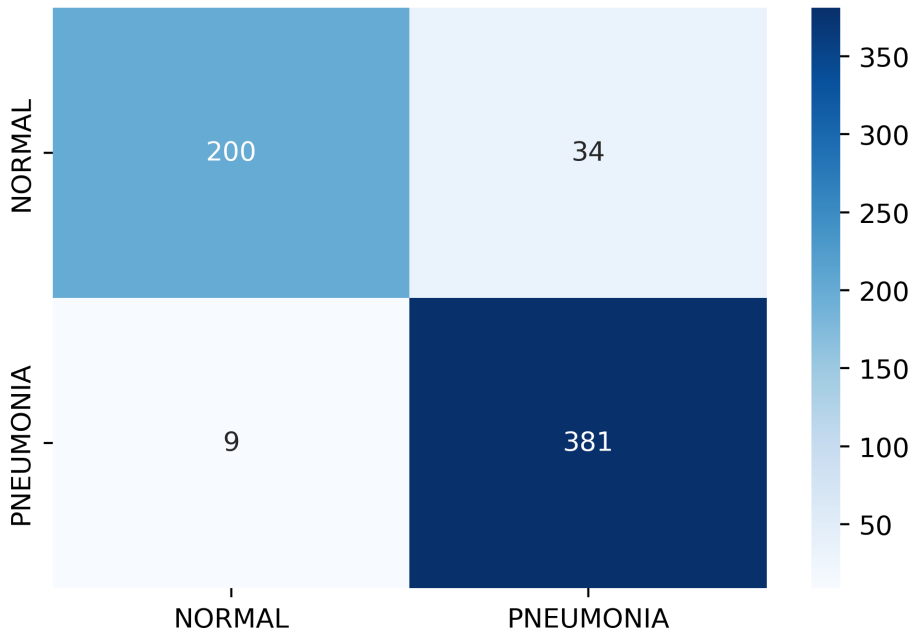
Loss



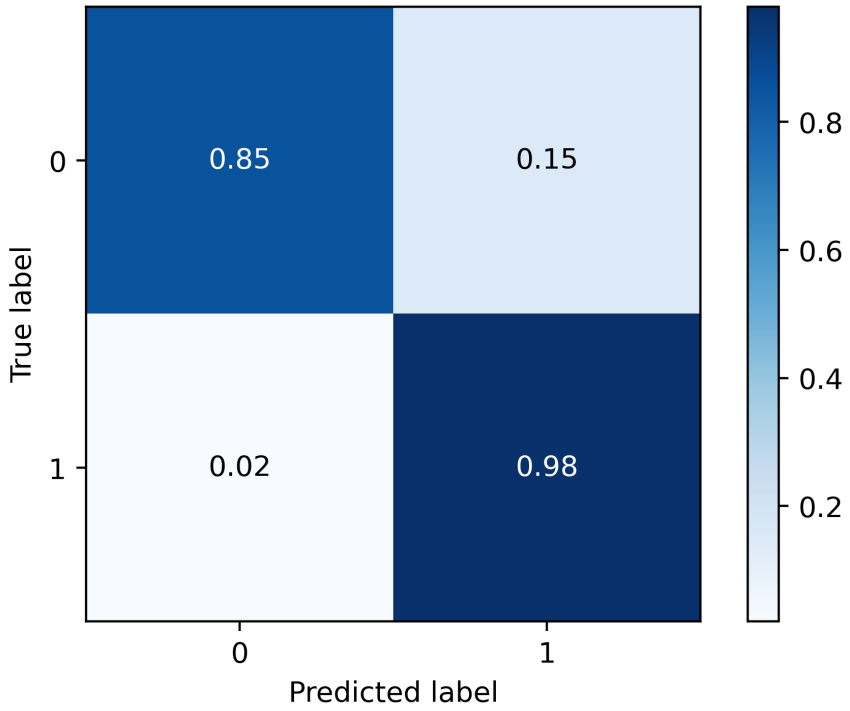
X_test

```
20/20 [==============================] - 54s 3s/step - loss: 0.2236 - accuracy: 0.9311
```

Confusion matrix



Normalized confusion matrix




```
Recall of the model is 0.98
Precision of the model is 0.92
```

Conclusiones

Transfer Learning

Diferencia significativa

VGG16

Generaliza muy bien acompañado de los parámetros correctos

Image Data Generator

Ayuda a aumentar el número de imágenes y mejora los resultados del modelo

Recursos

Chest X-Ray Images (Pneumonia)

kaggle

Desarrollo:

- Python
- Suite Anaconda
- Kaggle notebooks (GPU)

Librerías:

- Numpy
- Pandas
- Sklearn
- Tensorflow

Visualización:

- Matplotlib
- Seaborn
- Image.io
- Scikitplot

Slides:

- ALLppt

Información

Marcos Díaz Díaz

Bootcamp Data Science

The Bridge, Digital Talent Accelerator

Madrid, diciembre 2020

marcosdiak@gmail.com