

Proyecto 2:  
Predicción de edad ósea en niños

David Uriel Soto Álvarez 17551  
Rodrigo Samayoa Morales 17332  
María Fernanda Estrada Cornejo 14198  
Christopher Kevin Sandoval García 13660

Data Science 1



Universidad del Valle de Guatemala  
Facultad de Ingeniería  
Noviembre 2020

# Introducción

El estudio de la edad ósea permite a los médicos saber la madurez del sistema esquelético de un niño. El método más utilizado es mediante la radiografía de la mano izquierda, que va desde la muñeca hasta los dedos. Esta imagen es comparada con un atlas estándar del desarrollo óseo normal de niños de la misma edad y sexo. Este estudio normalmente lo solicita un pediatra para evaluar qué tan rápida o lenta es la maduración esquelética del niño, con el fin de determinar si padece de una enfermedad. El problema que tiene este método de comparación es que se pueden dar errores de interpretación, ya que depende absolutamente de la persona que compara. Además, si el niño se encuentra en la etapa temprana de una enfermedad ósea, esta comparación es muy importante. Por lo tanto, se pretende buscar una manera en la cual se pueda aplicar la tecnología y Data Science con el fin de automatizar este proceso y llegar a generar un modelo que permita hacer una clasificación confiable que permita poder dar diagnósticos más confiables, precisos y exactos para los pacientes. El siguiente informe presenta el proceso de determinar los mejores modelos para resolver este problema y cómo se elaboraron, sus resultados y un breve análisis comparativo.

## Objetivos

### General

El objetivo de este proyecto es poder realizar un análisis de imágenes de estructuras óseas, en particular del brazo, para poder automatizar este proceso que permite determinar la Edad Ósea (EO) del individuo.

### Específicos

1. Generar modelos y predicciones sobre la Edad Ósea con el uso de Data Science, a partir de técnicas de filtrado y procesamiento de imágenes.
2. Minimizar la cantidad de errores al realizar un diagnóstico de Edad Ósea, para que no se tenga que hacer un análisis subjetivo por parte de especialistas, el cual podría estar basado muchas veces en el sesgo de la persona.

# Marco teórico

## Análisis de radiografías desde el punto de vista médico

Este examen o análisis de radiografías es realizado por un médico, especialmente pediatras. El método más utilizado es mediante la radiografía de la mano izquierda, que va desde la muñeca hasta los dedos. Esta imagen es comparada con un atlas estándar del desarrollo óseo normal de niños de la misma edad y sexo. Es un procedimiento seguro e indoloro que utiliza una pequeña cantidad de radiación. La edad ósea se mide en años y meses. El análisis principal se da en el cartílago de las manos.

Este estudio puede ayudar a evaluar qué tan rápido o lento es el crecimiento del esqueleto del niño/niña, lo cual puede diagnosticar alguna enfermedad de crecimiento que padezca. La edad ósea también se puede utilizar para predecir por cuánto tiempo crecerá un niño, cuándo ingresará a la pubertad, cuál será la altura final del niño, o para controlar la evolución y guiar el tratamiento de los niños con afecciones que afectan el crecimiento.

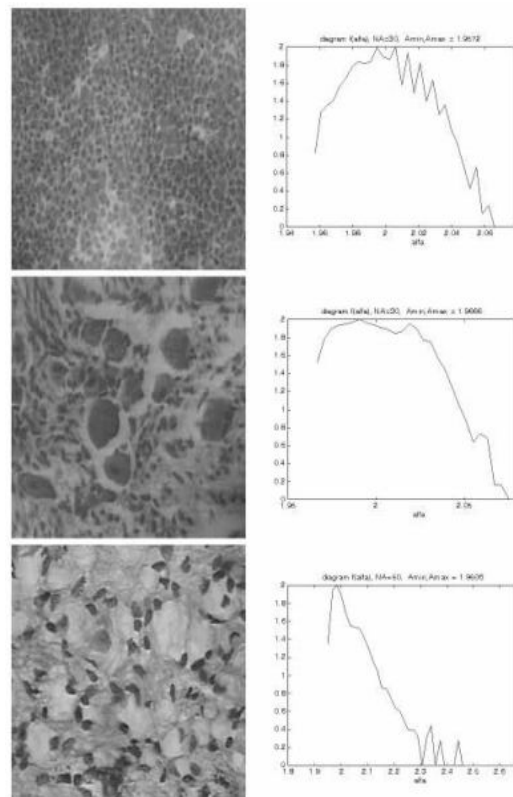
## Algoritmos de aprendizaje de máquinas y procesamiento de imágenes

- **Procesamiento de datos médicos mediante el análisis de imágenes y la geometría fractal**

Los avances topológicos han permitido que los objetos y sus propiedades sean analizados de maneras más complejas. Un ejemplo, es el cambio de perspectiva de análisis de superficies, pasando de Euclidiano a topológico. El análisis topológico permite descomponer el espacio a analizar en partes más pequeñas, para poder representarlo e incluso medirlo con más precisión. **Imagen 1. Análisis de imagen mediante MFA por el Instituto de Oncología de Yugoslavia**

De igual forma, se introdujo el análisis fractal de cualquier sistema complejo o irregular, permitiendo encontrar patrones y similitudes en donde aparentemente solo hay caos. Con estos nuevos conceptos introducidos en el área de la medicina, comenzaron a surgir investigaciones que relacionaban comportamientos bioquímicos luego de traumas, con multifractales. Fue así como el instituto de oncología de Yugoslavia introdujo el análisis multifractal (MFA) para clasificación de imágenes de 256x256 píxeles en escala de grises, para el análisis de traumas epidermicos por quemaduras.

Estas interpretaciones fueron útiles para el Instituto de Medicina Patológica y Forense de la



Academia Médica Militar de Belgrade, por lo que profundizaron el estudio y desarrollaron métodos y programas como FALFA Y HISTMF que seguirían mejorando y evolucionando hasta el día de hoy. (Reljin & Reljin, 2012)

- **Análisis de imágenes utilizando modelos de redes neuronales convolucionales**

Las redes neuronales convolucionales (CNN) es un algoritmo de Deep Learning utilizado principalmente para clasificar imágenes. Este algoritmo recibe como input una imagen, detecta las características principales de la misma y devuelve la categoría de la imagen. La forma en que detecta las características principales es que se le asigna un peso (importancia) a ciertos elementos de la imagen para poder diferenciar unos de otros. Algo importante a recalcar es que la salida se expresa en probabilidades, por lo que se debe tomar el que de mayor probabilidad.

El kernel en las CNN se considera como el filtro que se aplica a una imagen para extraer ciertas características o patrones importantes. Sirve para detectar bordes, enfoque, desenfoque, etc. Esto se logra al realizar una convolución entre el kernel y la imagen. La convolución en las redes neuronales consiste en tomar un grupo de píxeles de la imagen de entrada e ir realizando un producto escalar con un kernel. El kernel recorrerá todas las neuronas de entrada y obtendremos una nueva matriz, la cual se convierte en una de las hidden layers.

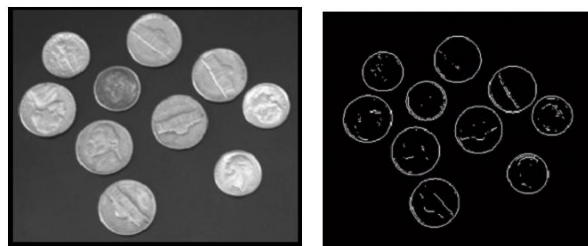


Imagen 2. Ejemplo de convolución entre el kernel y la imagen

Luego, en la capa de pooling, se combinan todas las salidas de los grupos de neuronas en una sola neurona para la siguiente capa. Esta capa es muy importante por el hecho que las imágenes no siempre se encuentran en la misma posición; pueden estar rotadas, posicionadas en la esquina, etc. Pooling reduce el tamaño de la representación de entrada y permite detectar objetos en una imagen sin importar dónde se encuentren. También ayuda a controlar el overfitting. Ya con estas capas creadas, se procede a la parte de flattening que consiste en aplanar estas matrices en un solo vector y ya el perceptrón hace la parte final de clasificación. (Bonner, 2019)

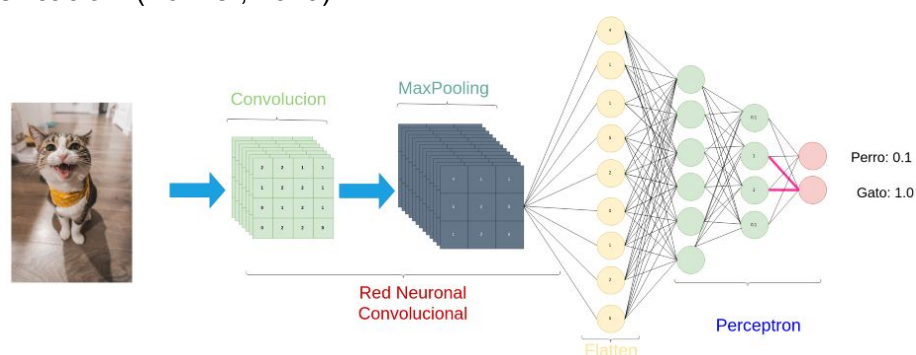
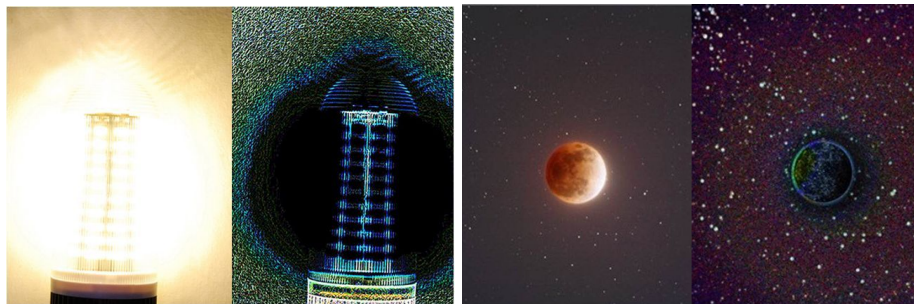


Imagen 3. Pasos de aplicación de una CNN

- **Refinamiento de imágenes con el algoritmo Phase Stretch Transform**

Existe una extensa cantidad de algoritmos que permiten el procesamiento de imágenes, los cuales permiten que se realicen cambios en ellas con el fin de poder mejorar su calidad, o en otros casos, permitir que la computadora o máquina pueda tener mayor facilidad al momento de reconocer o resaltar detalles. Estos algoritmos realizan acciones sobre las imágenes como por ejemplo cambiar la tonalidad, el contraste, mejorar definición, cambios de colores, rotación de contenido, reconstruir datos, entre otros. Estos algoritmos no proveen un tipo de modelo como tal para clasificación o predicción, pero son útiles al momento de apoyar a un algoritmo que sí haga este tipo de tareas facilitando el procesamiento o mejorando la calidad del modelo.

En este caso, el algoritmo Phase Stretch Transform es un algoritmo de procesamiento y reconocimiento de imágenes que es más o menos reciente, con 3 años de haber sido desarrollado con ingenieros de UCLA con acceso a código abierto. Este tiene como principal característica que mejora las imágenes, reconoce texturas y puede distinguir los bordes y otros detalles que aparecen en las imágenes. Capaz de mostrar detalles no visibles a simple vista, como estructuras de diversos materiales o las estrellas en una imagen del espacio. Algunos de los ejemplos que se han mostrado sobre cómo quedan las imágenes al usar el algoritmo son las siguientes:



**Imagen 4. Aplicaciones del algoritmo Phase Stretch Transform de los ingenieros de la UCLA sobre imágenes**

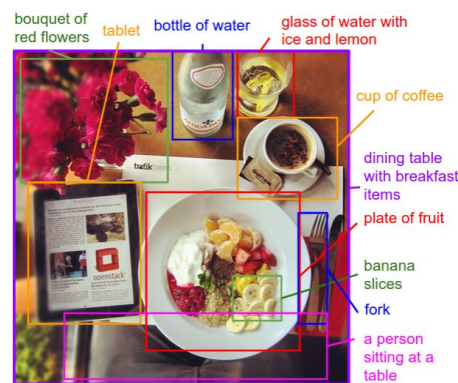
Otro dato interesante del algoritmo es que puede procesar imágenes a una alta velocidad, lo que puede llegar a permitir el análisis e identificación de imágenes de video, o bien el procesamiento de una gran cantidad de imágenes en poco tiempo. Otras aplicaciones del algoritmo que mencionan los investigadores de la UCLA son la detección de células cancerígenas, análisis de huellas dactilares, caras y ojos.

Una pequeña descripción sobre cómo funciona el algoritmo es que la imagen es pasada a través de un filtro de suavizado seguido de la aplicación de una fase dependiente de la frecuencia no lineal descrita por el kernel de la fase PST. La salida de la transformación es la fase en el dominio espacial. El paso principal es la función de fase 2-D (kernel de fase PST) que normalmente se aplica en el dominio de la frecuencia. La cantidad de fase aplicada a la imagen depende de la frecuencia con una mayor cantidad de fase aplicada a las características de frecuencia más alta de la imagen. Dado que las transiciones nítidas, como los bordes y las esquinas, contienen frecuencias más altas, PST enfatiza la información de los bordes (Chin, 2016).

- **Generar descripciones de imágenes con redes neuronales recurrentes**

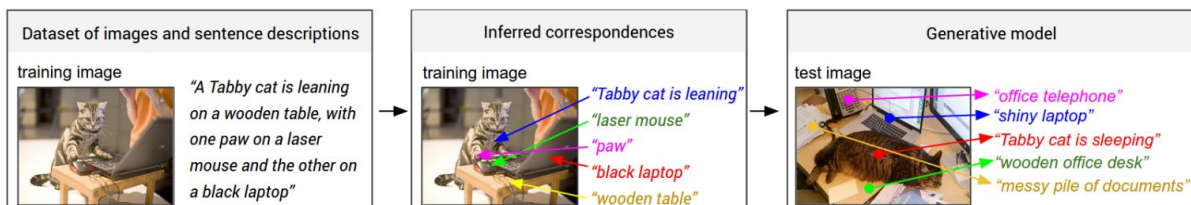
Se le conoce redes neuronales recurrentes a los algoritmos que implementan neuronas retroalimentadas, es decir neuronas que tienen capacidades de memoria en el tiempo. Para los humanos, es fácil describir una inmensa cantidad de detalles sobre una imagen. Sin embargo, esto es muy difícil realizarlo con modelos clásicos de reconocimiento visual. Normalmente el reconocimiento visual se ha centrado en etiquetar imágenes con un conjunto fijo de elementos visuales categorías. Sin embargo, mientras que los vocabularios cerrados de conceptos visuales nos otorgan modelos más sencillos de implementar, son muy restrictivos en comparación con los enormes cantidad de descripciones que un humano puede dar.

Para dar una descripción completa de una imagen utilizando redes neuronales recurrentes y convolucionales se analiza la imagen por regiones, identificando poco a poco el contexto de los posibles objetos que forman toda la imagen.



**Imagen 5. Identificación de contexto por sectores**

Luego basado en estos resultados se pueden unir las diferentes inferencias de la imagen en una descripción coherente y más completa de todo lo que se encuentra en la imagen original. El proceso de inferir a qué sectores corresponden diferentes descriptores se puede observar en la siguiente imagen. (Karpathy, *et. al* 2015)



**Imagen 6. Inferencia de descriptores**

- **Data Augmentation**

Un problema en Deep Learning es que para entrenar correctamente un modelo se necesita una gran cantidad de datos. En esto nos ayuda Data Augmentation, nos permite aumentar el dataset de imágenes de 2 formas: Introduciendo perturbaciones en las imágenes, así como replicando la imagen descentrada o invirtiendo los ejes y también utilizando distintas distribuciones, por ejemplo, si se quiere entrenar un modelo de clasificación de imágenes de alta resolución, se le añadirá imágenes con ruido, baja resolución, etc, pero siempre manteniendo la mayor proporción en las imágenes de alta resolución.

# Metodología

## Pasos para resolver el problema

Los siguientes pasos se tomaron para solucionar el problema planteado:

1. Realizar un análisis exploratorio de los datos y de las imágenes
2. Realizar limpieza de datos si era necesario
3. Investigar algoritmos de deep learning que ayudarían a resolver el problema.
4. Seleccionar algoritmos de deep learning y estudiarlos.
5. Preprocesar las imágenes.
6. Entrenar y generar modelos de predicción en Python.
7. Probar nuevos modelos con el dataset de prueba.
8. Crear visualizaciones estáticas y dinámicas, con el fin de dar a conocer la solución encontrada al problema.

## Selección de conjuntos de entrenamiento y prueba

De entrada, Kaggle proporciona ya separados los conjuntos de entrenamiento y de prueba. Sin embargo, se procuró que la división fuera más aleatoria. Entonces, se juntaron ambos conjuntos, se movieron aleatoriamente los datos, y se separó en un 70% de entrenamiento y 30% de prueba. De esta forma, la generación de modelos puede ser más precisa y tiene más variedad.

## Selección de algoritmos

- **Modelo de Convolución 2D - Keras**

Para este modelo se utilizó la librería Tensor Flow y Keras, donde Keras nos provee la facilidad de utilizar el algoritmo CNN para crear un red neuronal convolucional 2D. Se toma como entrada un arreglo tridimensional el cual es perfecto para el procesamiento de imágenes con 3 canales (RGB). Lo que hace es ir haciendo reducciones de matrices generando capas con las cuales se procede a generar un modelo que pueda ser ajustado durante el entrenamiento a modo de que al hacer una clasificación de una nueva imagen, ésta pueda ser evaluada en las distintas capas hasta obtener un resultado. En este caso sería la predicción de la edad ósea. En este caso sería de mucha utilidad para este tipo de clasificación/predicción de edad ósea y por eso lo utilizaremos.

- **Modelo Xception**

Al combinar las librerías de Matplotlib, Keras y Tensor Flow se cuenta con una gran variedad de funciones que permiten modificar el nivel de asertividad, según la capacidad de procesamiento y la magnitud de la data. Es por ello que se creó un modelo “entrenable” en el que se puede ajustar la cantidad de nodos y filtros a utilizar sobre la data. De esta forma no solo se toman en cuenta las dimensiones de cada elemento, sino también la cantidad de conexiones que se harán entre ellos. De la misma manera, al trabajar en Google Colab, se

tienen ciertas facilidades de código. La mayor ventaja fue trabajar con un entorno GPU que permite una mayor velocidad de procesamiento que en el servidor local.

- **Modelo VGG 16**

El modelo VGG 16 es una rama de las redes neuronales convolucionales. Una gran ventaja de este modelo es el tamaño pequeño del kernel, ya que sigue teniendo la misma eficiencia pero con un kernel menor a otros modelos. Sin embargo, tiene el problema de llevar más tiempo en entrenar y tiene pesos grandes; funciona mejor con set de datos pequeños. A pesar de esto, se seleccionó este modelo por ser simple de entender e implementar. Es parte de Keras, por lo que la documentación es fácil de encontrar.

## **Selección de herramientas utilizadas**

Al inicio se seleccionó Kaggle para correr los jupyter notebooks, ya que proporciona un ambiente virtual en la nube. Sin embargo, al tratar de guardar el modelo ya entrenado para ingresarlo a la visualización gráfica, las versiones de h5 eran incompatibles. Entonces cada modelo fue entrenado en la computadora de integrantes designados para tener compatibilidad. Para ello fue necesario que uno de los recursos de cómputo fuera una RAM potente.

El lenguaje para generar los modelos y la visualización interactiva fue Python, ya que es un lenguaje sencillo de utilizar y se especializa en Deep Learning. La biblioteca principal utilizada fue Keras. Keras proporciona implementación de los tres modelos seleccionados.



# Resultados y análisis de resultados

## Características del conjunto de datos original

El set de datos a analizar consiste en 12800 imágenes de radiografías de mano, cada una acompañada de su id, género y edad ósea en meses.

- Variables

Nombre	Categoría	Descripción	Ejemplo
id	Categoría nominal	Indica a qué imagen están relacionados los datos de género y edad ósea. Solamente es un identificador.	rango: 1377 - 15.6k 1377, 1378, 1379...
boneage	Numérica discreta	Indica la edad ósea en meses de la imagen.	rango: 1 - 288 180, 12, 94, 120...
male	Categoría binaria	Indica si la radiografía es de sexo masculino o no. Es un valor booleano; si es True es masculino, si es False es femenino.	True, False

## Tareas de limpieza y procesamiento

El set de datos es principalmente numérico, por lo que no se necesitaron muchas operaciones de limpieza. Sin embargo, para simplificar aún más, se cambiaron los valores de la columna male. Se realizó lo siguiente:

1. Verificar que solo existan datos numéricos en las columnas de id y boneage.
2. Verificar que solo existan valores de True y False en la columna de male.
3. Cambiar los valores de True y False a 1 y 0, respectivamente en la columna male.
4. Verificar que hay la misma cantidad de imágenes y datos relacionados.

Por otro lado, a todas las imágenes se les realizó un procesamiento en cada modelo para incrementar el porcentaje de accuracy. Algunas correcciones que se hicieron, en general, fueron: recortar las imágenes, incrementar el contraste, cambiar a escala de grises, etc.

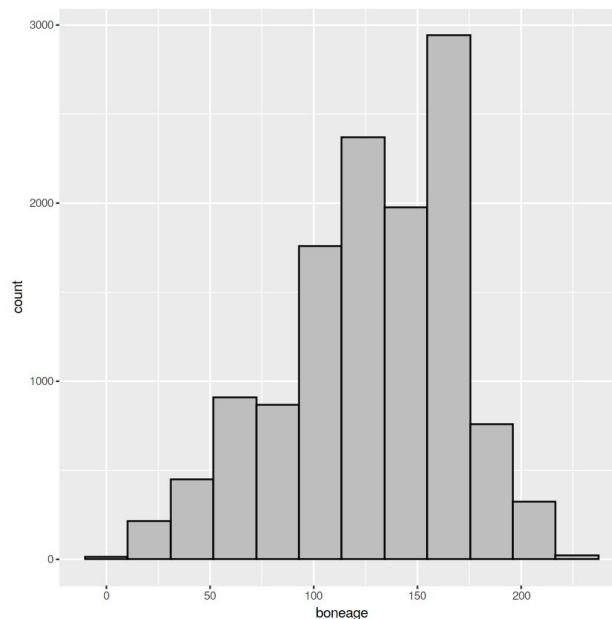
## Análisis exploratorio

- Variable cuantitativa: boneage
  - Resumen

```
boneage
Min.   : 1.0
1st Qu.: 96.0
Median :132.0
Mean   :127.3
3rd Qu.:156.0
Max.   :228.0
```

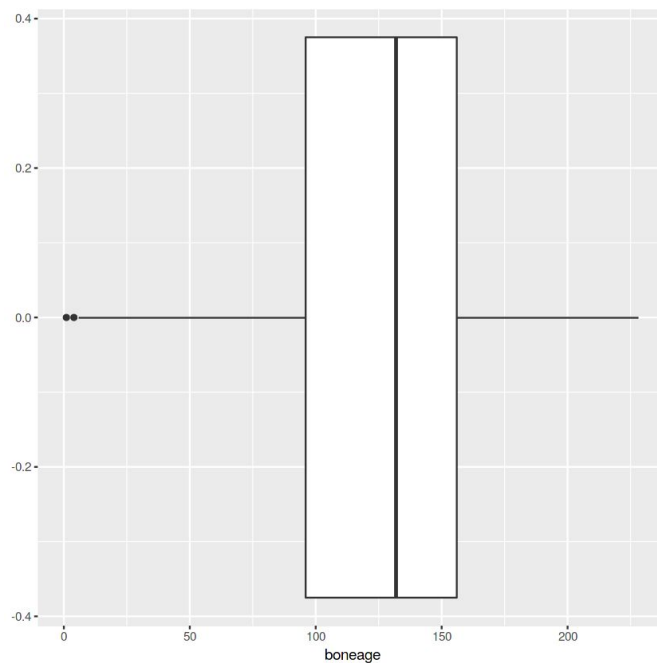
Como se puede observar en la tabla de resumen que se realizó para la variables “bonage”, se puede ver que esta tiene un rango de valores entre 1 a 228, lo cual vendría siendo la edad ósea en años en que está acotada el dataset. También es posible observar que el primer cuartil de los datos está entre el dato en los 96 años, que el segundo cuartil o la mediana está a los 132 años, y por último el tercer cuartil está a los 156 años. Esto podría darnos una idea de que el primer cuartil o las edades óseas más bajas son las que tienen mayor presencia o variedad en este dataset. Por otro lado es interesante ver que la media, que es 127 años es muy similar a la media por lo que tal vez podríamos tener una distribución normal de la muestra de datos, o quizás existan algunos datos atípicos que generan este resumen estadístico.

- Histograma



Se puede observar una distribución casi normal en los datos de la variable boneage, con media en 127 años, tal y como se observó en el resumen de los datos. No es completamente normal debido a que existe un pequeño sesgo a la derecha, pero aun así este sesgo es parte de los datos centrales, por lo cual se podría decir que los datos están bastante normalizados. Cabe destacar que las edades “extremos” son poco frecuentes y que mientras más nos acercamos a la media, más casos o datos se tienen.

- Diagrama de caja y bigotes



En esta gráfica podemos observar que no hay muchos datos atípicos en la variable de boneage y que esta variable tiene un rango intercuartil aproximadamente entre 100 y 150. También podemos observar que hay un ligero sesgo a la derecha en los datos el cual es normal dentro de lo que se puede esperar en un dataset que tiene una distribución casi normal.

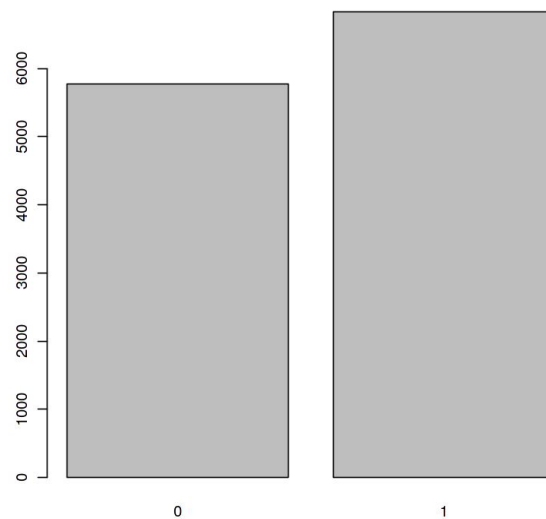
- Variable cualitativa: male
  - Tabla de frecuencia

```
male
False:5778
True :6833
```

Se puede observar en la tabla de frecuencia de la variable “male”, que existen únicamente dos posibles valores: True y False. True indica que la radiografía pertenece a un hombre y False indica que pertenece a una mujer. Como indica la tabla, en el set de datos hay más muestras de hombres que de mujeres, con una diferencia de 1055 datos. A pesar de que es una diferencia amplia en los datos, dado que se está trabajando en el orden de los miles, cabe dentro de lo normal tener una diferencia de este tipo, y que los datos sean útiles para el estudio.

Dado que el texto que contiene la variable es poco útil y es una categoría binomial, entonces se decidió seguir analizando esta variable haciendo la sustitución de True y False por 1 y 0 respectivamente.

- Gráfico de barra - Donde 0 es False y 1 es True



De una forma más gráfica, se muestra que hay más cantidad de hombres (variable 1) que de mujeres (variable 0). Sin embargo, considerando la cantidad de datos de cada uno, la diferencia es mínima. Por lo mismo que se mencionó anteriormente, se determinó que el conjunto de datos está lo bastante normalizado y distribuido para realizar un estudio en donde se pueda considerar tanto a hombre como mujeres.

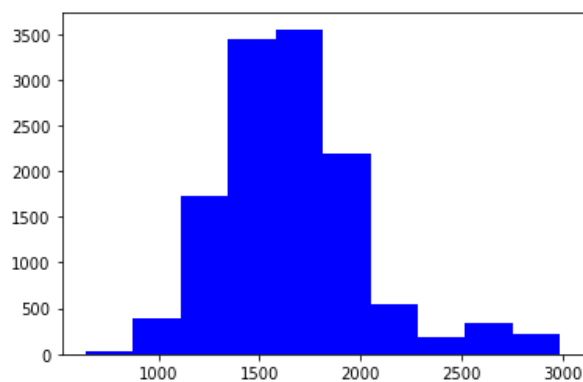
- Tabla de proporción - Donde 0 es False y 1 es True

	0	1
	0.4581714	0.5418286

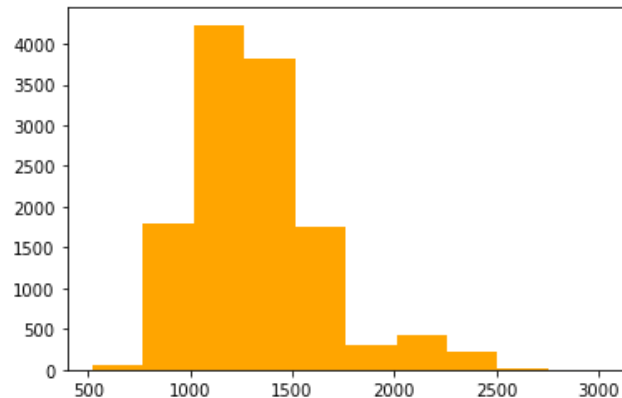
En esta tabla de proporción es más notorio que la diferencia entre la cantidad de datos para hombres y mujeres es mínima, ya que la proporción de ambos es cercana al 50%. Con esto en mente, podríamos inferir que la modelación que se hará a partir de estos datos para poder clasificar y dar diagnósticos de edad ósea, será más probable que no se tengan sesgos causados por el sexo de la persona, ya que ambos sexos están siendo tomados en cuenta en proporciones casi iguales.

- Imágenes de radiografías:

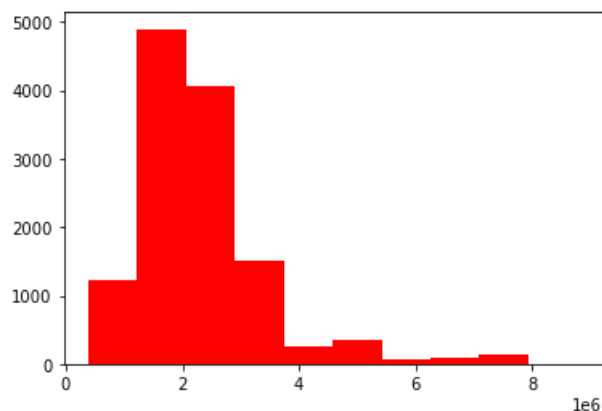
- Histogramas



Este histograma muestra el rango de valores de height del set de imagenes proporcionado. La altura más grande es de 2989 pixeles y la más pequeña de 640 pixeles. También se puede observar que la mayoría de imágenes tienen altura entre 1200-2000 pixeles.



Este histograma muestra el rango de valores de width del set de imagenes proporcionado. La anchura más grande es de 3001 pixeles y la más pequeña de 521 píxeles. También se puede observar que la mayoría de imágenes tienen anchura entre 800-1600 pixeles.



Este histograma muestra el rango de valores de píxeles del set de imagenes proporcionado. La cantidad de píxeles más grande es de 8765921 y la más pequeña de 387103. También se puede observar que la mayoría de imágenes tienen cantidad de pixeles entre 0.5-3.5 millones.

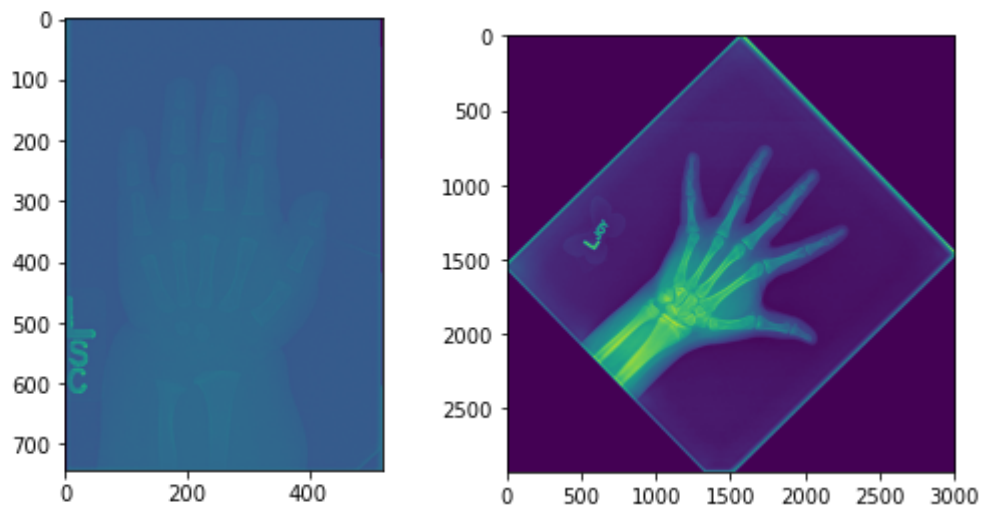
- Resumen

A fin de tener recopilados los datos anteriormente mencionados, se generó un resumen como sigue

```
La altura mas grande es: 2989
La altura mas pequeña es: 640
La anchura mas grande es: 3001
La anchura mas pequeña es: 521
La cantidad de pixeles mas grande en un imagen es: 8765921
La cantidad de pixeles mas pequeña en un imagen es: 387103
```

- Ejemplos de radiografías

Ya con los datos de cuál es la altura y anchura más grande y más pequeña, se procedió a identificar la diferencia entre ambas imágenes. Se observa que la radiografía más pequeña está muy borrosa y no se distinguen bien los huesos de la mano. A diferencia de esta, la radiografía más grande muestra gran detalle de los huesos. Es importante notar que la posición de las manos en ambos casos es totalmente distinta, por lo que se debe considerar rotar y recortar ciertas imágenes. A continuación se muestran ambas radiografías.



## Ajuste de parámetros a modelos

A los tres modelos se les modificaron los mismos parámetros. Como base, se mejoró el procesamiento de las imágenes para incrementar el porcentaje de accuracy. Luego, se modificaron las iteraciones que se establecieron inicialmente, para ser aproximadamente 15-20 iteraciones. Después, se ajustó la cantidad de meses de la comparación por rango, quedando el primer modelo con 30 meses y el resto con 15. Por último, se modificaron valores específicos como el tamaño de kernel, número de capas y filtros, etc. Esto fue específico de cada modelo y dependía del resultado del accuracy final.

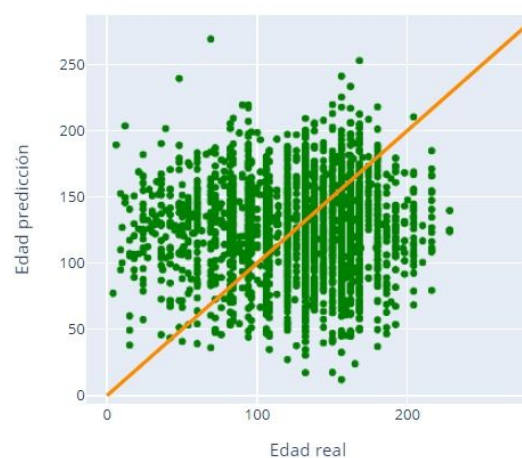
## Comparación de algoritmos

- **Modelo de Convolución 2D - Keras**

En este caso el modelo no fue tan eficiente. Esto se determinó ya que al generar el modelo se utilizó el indicador del MSE (Error medio cuadrático) para poder obtener resultados sobre qué tan alejado del resultado estaba el valor real mientras el modelo se entrenaba. Y este comenzó con un error medio de aproximadamente 1300 meses, lo cual era demasiado, pero con 15 iteraciones que fueron las que se establecieron para generar el modelo para entrenar, este se redujo a alrededor de 250 meses. Esto aún es bastante, pero tomando en cuenta que eran muchos datos, quizás al sacar el error medio, el problema estaba más presente en algunas imágenes que en otras.

Esto debido a que cuando se evaluaron los resultados de algunas predicciones se podía observar que la predicción fallaba por 30 meses en algunas, en otras 70 meses, y en unas 100 meses o más. Pero en su mayoría el error era de unos 30 a 70 meses. Probablemente, algo que pudo mejorar pudieron ser los filtros a las imágenes para que el algoritmo pudiera identificar y clasificar de mejor manera en las 15 iteraciones. Lo positivo, fue observar que los resultados obtenidos iban acorde a edades óseas y no generación de predicciones con valores completamente erróneos. Dado que el resultado era una predicción de un número decimal que podía variar y no una etiqueta fija, es comprensible que el algoritmo no sea tan acertado. Haciendo una comparación por rangos de 30 meses se obtuvo un accuracy del 48.6%, por lo cual no es un modelo tan acertado como se esperaba.

Convolución 2D - Keras: 48.83% Accuracy

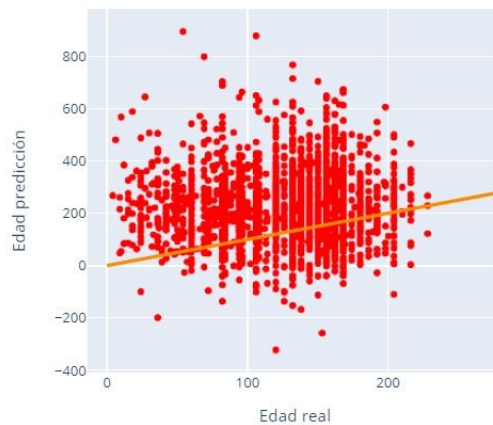


- **Modelo VGG 16**

Este modelo logra clasificar categorías específicas y no números como lo es la edad ósea. Es por esto que antes de entrenar al modelo, se determinaron 10 grupos para las edades óseas (de 1-228 con saltos de 22) para que cada imagen se clasificará en alguno de estos grupos. El resultado inicial con dos capas, kernels de 5x5 y filtros de 20 y 50 respectivamente, dieron como resultado un accuracy del 9% después de 10 iteraciones. Inicialmente se determinó que el preprocesamiento de las imágenes no fue el adecuado, por lo que se aplicaron más modificaciones a las que no fueran tan claras o que se encontraran en una posición extraña. Sin embargo, después de hacer las nuevas modificaciones a las imágenes, el accuracy aumentó sólo un 5.2%.

Para mejorar el accuracy obtenido, se tomó la alternativa de modificar los valores de kernel, cantidad de filtros y número de iteraciones. Después de estos cambios, el accuracy llegó hasta 16.57% con 20 iteraciones. Un resultado similar se obtuvo al aplicar el modelo en el dataset de test. Al igual que en otros modelos, se pudieron mejorar los filtros de las imágenes, utilizando blanco y negro o aumentando más el contraste. Por otro lado, este modelo sólo clasificaba las imágenes en alguno de los 10 rangos definidos, por lo que no se puede dar un valor exacto de la edad ósea, solo un aproximado. Aunque entrenar el modelo tardó poco, se vio afectado el accuracy significativamente.

VGG16 - Keras: 16.57% Accuracy

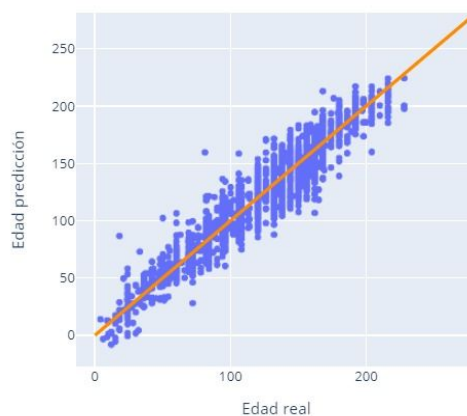


- **Modelo Xception**

Se inició importando todos los paquetes de Matplotlib, Keras y Tensor Flow necesarios para poder diseñar el modelo de predicción. Luego se definieron 20 epochs debido a que daría un nivel de asertividad aceptable sin recargar el peso de procesamiento. De igual forma, se estableció que cada elemento tendrá 3 dimensiones y se hizo un análisis de dimensiones para normalizar cada imagen. Al normalizarlas se lograría quitar cualquier tipo de rotación o zoom que se pudiese tener.

Al diseñar el modelo de predicción, se utilizó un formato secuencial y para maximizar el nivel de asertividad se utilizaron las funciones flatten y dense, mediante la función Max Pooling. Obteniendo una cantidad de parámetros entrenables de 22,431,617. A partir del modelo realizado, se tomó como indicador de efectividad el error absoluto medio. Esto dio un resultado de alrededor de 15 meses, el cual mejoró significativamente con respecto a los otros 2 modelos. Al hacer un análisis de rango de 15 meses se obtuvo un accuracy de 77.45%. El entrenamiento tardó más que los otros dos modelos, pero el resultado fue más satisfactorio.

Xception - Keras: 77.45% Accuracy





## Aplicación creada

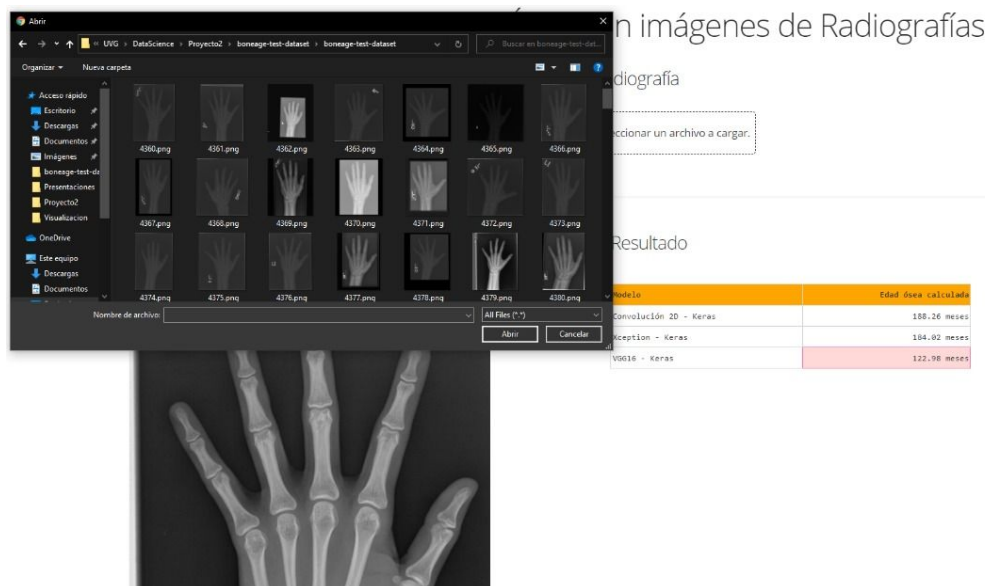
Como todos los modelos generados fueron con el lenguaje Python, se pretendió investigar y utilizar una herramienta visual en Python. Es por esto que se seleccionó el framework “dash”, el cual está diseñado especialmente para Machine Learning y Data Science web apps. Incluye dropdowns, sliders, gráficas, etc., por lo que es muy intuitiva.

Al iniciar la app, se presenta una pantalla con información general de lo que el usuario puede hacer. Como cargar imágenes de prueba para predecir y obtener el resultado de cada modelo, generar los plots de cada modelo y mostrar los porcentajes de accuracy.

Pantalla inicial:



Cargar radiografía y predecir:



Radiografía

1377.png



Resultado

Modelo	Edad bones calculada
Convolution 2D - Keras	180.26 meses
Xception - Keras	184.02 meses
VGG16 - Keras	122.98 meses



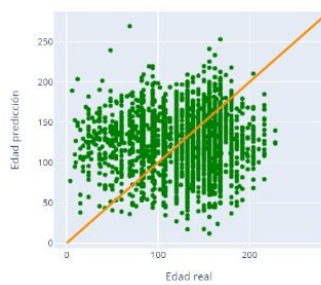
Mostrar modelos y su porcentaje de accuracy:

Mostrar el rendimiento de los algoritmos

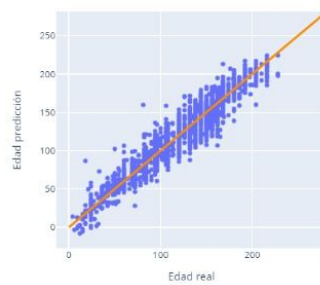


Rendimiento de los algoritmos con datos de prueba

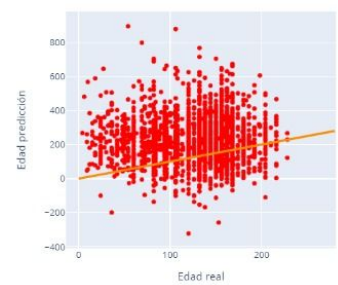
Convolución 2D - Keras: 48.83% Accuracy



Xception - Keras: 77.45% Accuracy



VGG16 - Keras: 16.57% Accuracy



## Conclusiones

Con respecto al modelo de Convolución 2D se realizaron un máximo de 15 iteraciones debido a que el modelo era muy tardado para optimizar el resultado. Sin embargo, el mejor resultado que se pudo obtener fue tener un error medio cuadrático (MSE) de 250 meses. Para poder obtener un resultado de porcentaje de efectividad del modelo se permitió que el resultado tuviera una variabilidad de 30 meses, con lo que el modelo obtuvo una efectividad del 48.6% al utilizar el set de imágenes para validación.

Por otro lado, el modelo de VGG 16 obtuvo un porcentaje de efectividad del 16.57% al realizar 20 iteraciones para entrenar el modelo. Este modelo fue consistente al utilizar el set de imágenes para validación. Al compararlo con el primer modelo se puede observar claramente que este tuvo un descenso considerable, principalmente por el tiempo de ejecución y su complejidad. Las ventajas de este modelo es que trabaja un poco más rápido para el proceso de entrenamiento y predicción, pero a costa de su accuracy.

Al utilizar los paquetes de tensor flow y keras, se tiene la facilidad de modificar la cantidad de filtros y nodos que se le apliquen a los elementos de nuestro set. Así se puede manipular el grado de accuracy que se desea, tomando en cuenta que el tiempo de procesamiento será mayor o menor dependiendo de ello. De igual forma, se pudo seleccionar la cantidad de epochs necesarios para mejorar la asertividad y según el modelo que se desee utilizar. Cabe mencionar que se utilizó un entorno GPU para poder trabajar con la magnitud del set en un tiempo considerable. Este último modelo mejoró considerablemente, obteniendo un porcentaje de efectividad del 77.45%.

Con lo discutido anteriormente, se determinó que el mejor modelo de predicción para la edad ósea a través de imágenes es el modelo de Xception, tanto por su rendimiento al momento de entrenar y predecir, y porque es el modelo que obtuvo el mejor porcentaje de efectividad.

Con los modelos generados, se concluye que el objetivo principal fue cumplido al automatizar el proceso de determinar la edad ósea del individuo. De igual manera el segundo objetivo se cumplió al haber investigado y generado tres modelos de Deep Learning. Por último, con el segundo modelo de Xception se cumplió el tercer objetivo al minimizar la cantidad de errores al realizar un diagnóstico.

## Referencias y bibliografía

Chin, M. (10 de febrero de 2016). *UCLA researchers release open source code for powerful image detection algorithm*. Science and Technology. Extraído de: <https://newsroom.ucla.edu/releases/ucla-researchers-release-open-source-code-for-powerful-image-detection-algorithm>

Gandhi, A. (2018). *Data Augmentation | How to use Deep Learning when you have Limited Data — Part 2*. Extraído de: <https://nanonets.com/blog/data-augmentation-how-to-use-deep-learning-when-you-have-limited-data-part-2/>

Reljin, I., & Reljin, B. (4 de octubre de 2012). *Fractal Geometry and Multifractal in Analyzing and Processing Medical Data and Images*. Belgrade, Yugoslavia: Facultad de Ingeniería.

Bonner, A. (2 de febrero del 2019). *The Complete Beginner's Guide to Deep Learning: Convolutional Neural Networks and Image Classification*. Extraído de: <https://towardsdatascience.com/wtf-is-image-classification-8e78a8235acb>

Karpathy, A. & Fei-Fei, L. (2015). *Deep Visual-Semantic Alignments for Generating Image Descriptions*. Extraído de: [https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2015/papers/Karpathy\\_Deep\\_Visual-Semantic\\_Alignments\\_2015\\_CVPR\\_paper.pdf](https://www.cv-foundation.org/openaccess/content_cvpr_2015/papers/Karpathy_Deep_Visual-Semantic_Alignments_2015_CVPR_paper.pdf)

Durani, Y. s.f. *Radiografía: estudio de la edad ósea*. Consultado el 07/09/2020 de <https://kidshealth.org/es/parents/xray-bone-age-esp.html#:~:text=Qu%C3%A9%20es,una%20peque%C3%B1a%20cantidad%20de%20radiaci%C3%B3n>.

Navarro, M.; Tejedor, B.; López, J. 2014. *El uso de la edad ósea en la práctica clínica*. Consultado el 07/09/2020 de <https://www.elsevier.es/es-revista-anales-pediatria-continuada-51-articulo-el-uso-edad-osea-practica-S1696281814702045>

Pérez, R. 2011. *Valoración y utilidad de la edad ósea en la práctica clínica*. Consultado el 07/09/2020 de <https://fapap.es/articulo/180/valoracion-y-utilidad-de-la-edad-osea-en-la-practica-clinica>