

A screenshot of a computer

Description automatically generated

**Situación Problemática**

Según la OMS, las enfermedades cerebrovasculares -ECV- son fenómenos agudos que se deben principalmente a obstrucciones que imposibilitan la circulación de la sangre hacia el corazón y el cerebro (Organización Mundial de la Salud -OMS-, 2017). Las ECV son la tercera causa de muerte, la segunda causa de discapacidad en adulto y la segunda causa de demencia en el mundo (Bleder, 2019).

Se estima que alrededor de 15 millones de personas sufren un ictus (derrame cerebral, embolia, trombosis o apoplejía) anualmente en el mundo, de las cuales 5.5 millones mueren y 5 millones más quedan con alguna discapacidad permanente. Así, en el mundo la incidencia promedio de las enfermedades cerebrovasculares es de 200 casos por cada 100.000 habitantes anualmente (Blender, 2019).

Según MayoClinic, las enfermedades cerebrovasculares pueden ser isquémicas o hemorrágicas, de acuerdo con sus características fisiopatológicas:

* Accidente cerebrovascular Isquémico: Este es el tipo más común de accidente cerebrovascular. Ocurre cuando los vasos sanguíneos del cerebro se estrechan o se bloquean, lo que causa una importante reducción del flujo sanguíneo (isquemia). Los vasos sanguíneos se bloquean o se estrechan debido a la acumulación de depósitos de grasa o de coágulos sanguíneos u otros desechos que se desplazan por la sangre (con más frecuencia desde el corazón) y se alojan en los vasos sanguíneos del cerebro. Algunas investigaciones iniciales muestran que la infección por COVID-19 puede aumentar el riesgo de sufrir un accidente cerebrovascular isquémico, pero se requieren más estudios.
* Accidente cerebrovascular hemorrágico: El accidente cerebrovascular hemorrágico ocurre cuando un vaso sanguíneo en el cerebro gotea o se rompe. Las hemorragias cerebrales pueden ser el resultado de muchas afecciones que afectan los vasos sanguíneos. Los factores relacionados con el accidente cerebrovascular hemorrágico incluyen los siguientes:
  + Presión arterial alta no controlada
  + Sobretratamiento con anticoagulantes
  + Bultos en puntos débiles de las paredes de los vasos sanguíneos (aneurismas)
  + Traumatismo (como un accidente automovilístico)
  + Depósitos de proteína en las paredes de los vasos sanguíneos que provocan debilidad en la pared del vaso (angiopatía amiloide cerebral)
  + Accidente cerebrovascular isquémico que causa hemorragia

**Problema científico**

Luego del primer ictus, puede ocurrir uno segundo empeorando las posibilidades de supervivencia del paciente, sin embargo, esto podría mitigarse si los médicos pudieran determinar la etiología del accidente cerebrovascular, lo que influye en el manejo terapéutico después de un accidente cerebrovascular (Kaggle, 2022).

Durante la última década, la trombectomía mecánica se ha convertido en el tratamiento estándar para el accidente cerebrovascular isquémico agudo por oclusión de grandes vasos. Como resultado, los coágulos recuperados se volvieron susceptibles de análisis. Se han intentado aplicar métodos basados ​​en el aprendizaje profundo para predecir la etiología del accidente cerebrovascular isquémico y el origen del coágulo. Lamentablemente, los formatos de datos únicos, los tamaños de archivos de imágenes y la cantidad de diapositivas de patología disponibles han dificultado solucionar el problema (Kaggle, 2022).

Para disminuir las posibilidades de sufrir accidentes cerebrovasculares posteriores, se debe desarrollar un modelo de inteligencia artificial altamente preciso y efectivo que pueda diferenciar de manera confiable entre las dos principales etiologías de los accidentes cerebrovasculares isquémicos agudos (AIS): la causa cardíaca y la aterosclerosis de arteria grande, utilizando imágenes completas de patología digital. Esto podría tener un gran impacto para salvar las vidas de los sobrevivientes de accidentes cerebrovasculares y ayudar a comprender mejor la segunda causa de muerte en el mundo.

**Objetivos**

**Objetivo general**

Desarrollar un modelo de inteligencia artificial efectivo y preciso que pueda distinguir entre las dos principales etiologías de los accidentes cerebrovasculares isquémicos agudos (AIS) - causa cardíaca y aterosclerosis de arteria grande - utilizando imágenes completas de patología digital, con el propósito de mejorar la identificación de los orígenes de los coágulos sanguíneos en accidentes cerebrovasculares mortales y facilitar la prescripción de un tratamiento terapéutico post-AIS más efectivo, reduciendo así la probabilidad de un segundo AIS

**Objetivos específicos**

* Crear un conjunto de algoritmos de procesamiento de imágenes que permitan la extracción y caracterización de las características patológicas relevantes en las imágenes completas de patología digital relacionadas con los coágulos sanguíneos en AIS de origen cardíaco y aterosclerosis de arteria grande.
* Entrenar y evaluar un modelo de inteligencia artificial utilizando el conjunto de datos de imágenes completas de patología digital, con el objetivo de alcanzar una precisión mínima del 95% en la clasificación de la etiología de los coágulos sanguíneos en AIS, diferenciando con alta sensibilidad y especificidad entre las causas cardíacas y la aterosclerosis de arteria grande.

**Descripción de los datos**

Se cuenta con 3 carpetas con imágenes distribuidas de la siguiente manera

* Train: imágenes a utilizar para el proceso de entrenamiento en formato TIFF.
* Test: 280 imágenes comprimidas para utilizar en el proceso de testeo.
* Other: suplement de imágenes con una etiología distinta a CE o LAA o desconocida.

Además, de 3 archivos en csv con la siguiente información

* Train: contiene 5 variables con 754 observaciones:
  + Image\_id: identificador único que se forma como{patient\_id}\_{image\_num} y corresponde a la imagen {image\_id}.tif.
  + Center\_id: identifcador del centro médico donde se obtuvo la slide.
  + Patient\_id: identifcador del paciente de quien se obtuvo la slide.
  + Image\_num: enumera las imágenes de los coágulos obtenidos por el mismo paciente.
  + Label: etiología del coágulo, es decir, si es CE o LAA.
* Test: contiene 4 variables, idénticas a las de Train, a excepción del label con 4 observaciones.
* Other: contiene 5 variables, idénticas a la de Train a excepción de Center\_id que es reemplazada por other\_specificied que, en caso se desconozca la etiología, se etiqueta como “otro”, cuenta con 396 observaciones

**Análisis exploratorio**

La base de datos tiene las siguientes cinco variables, cuyos tipos son:

* image\_id: object
* center\_id: int64
* patient\_id: object
* image\_num: int64
* label: object

Tablas de frecuencia:

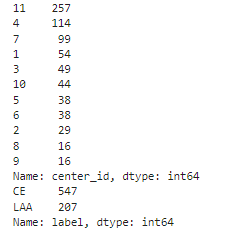
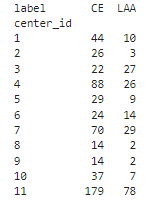
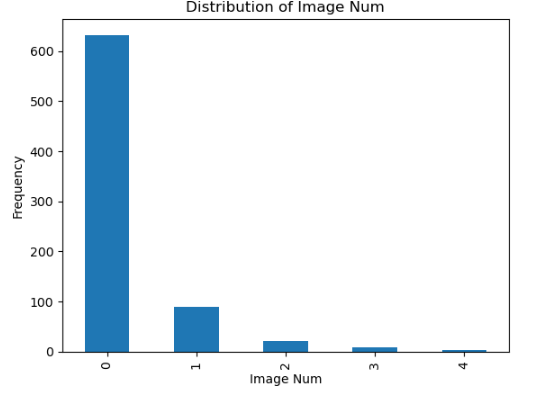


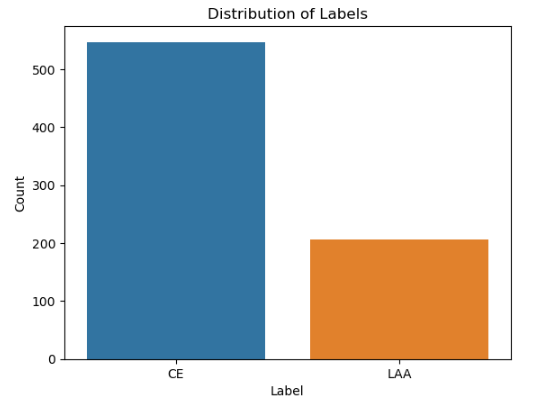
Tabla de contingencia:



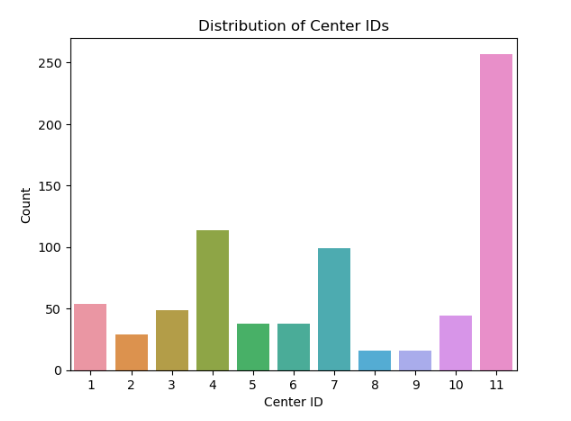
Primero estudiamos como se distribuyen las variables:



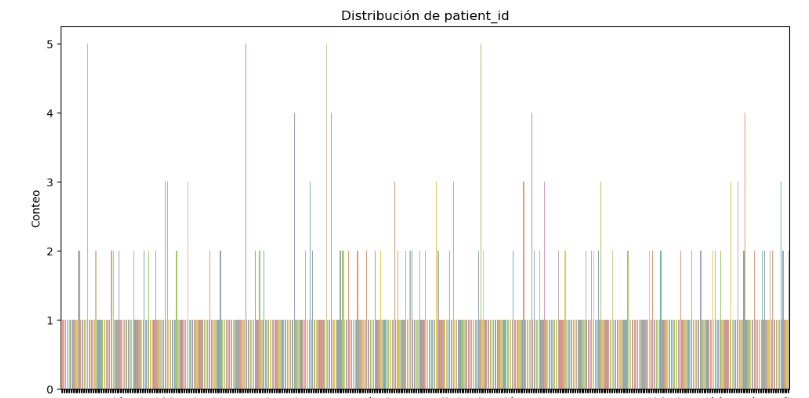
Se puede observar que la mayoría de los datos de la variable image\_num son 0.



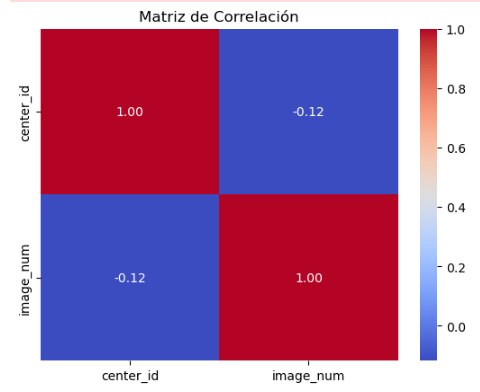
Vemos también que la mayoría de los datos tienen valor de label “CE”.



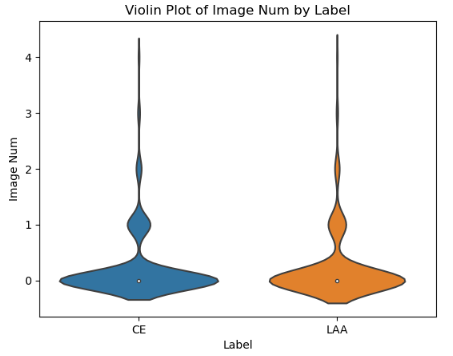
El centro que mas muestras contiene es el 11, mintras que los que menos tienen son el 8 y 9.



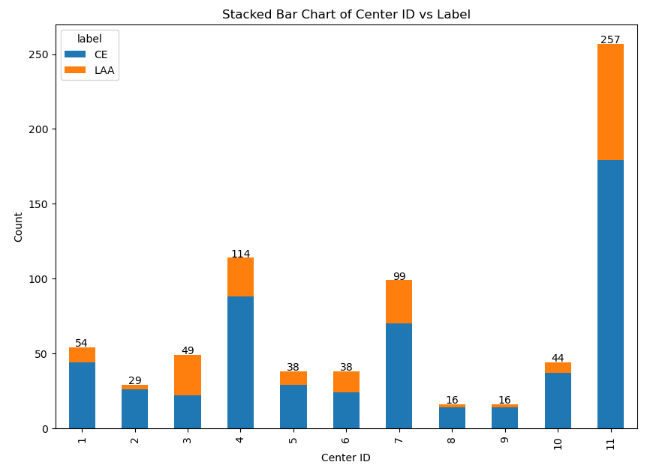
Se puede observar también que la gran mayoría de pacientes solo están una vez, sin embargo, hay algunos que aparecen más, llegando a haber pacientes con un máximo de cinco imágenes en el dataset.



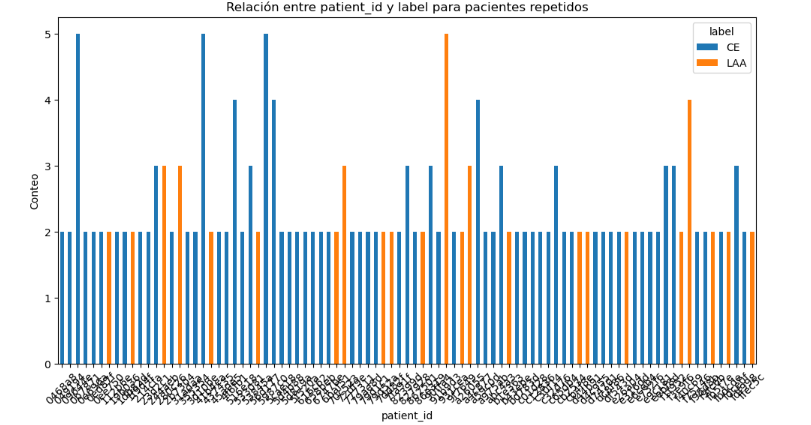
Al cruzar las variables vemos que no hay una correlación entre la variable image\_num y la variable center\_id, que son nuestras variables numéricas.



Entre image\_num y label vemos que en ambos valores de label se puede ver que los valores de image\_num tienen comportamientos muy similares, concentrando en ambos casos, la mayoría en el 0.



Entre las variables center\_id y label, vemos que en cada centro predomina la etiqueta “CEE”, y que en cada centro se encuentran ambas labels.



Buscamos también analizar si hay ciertas tendencias entre los tipos de coágulos de los pacientes que aparecen más de una vez en el dataset.

**Hallazgos y conclusiones**

* Los coágulos con etiología cardioembólica (CEE) son más frecuentes que los coágulos con etiología aterosteclórica (LAA).
* Los pacientes que han presentado coágulos en más de una ocasión, suelen presentar siempre el mismo las veces que se repiten. No se encontró ningún caso en que se presente uno y luego otro.

Bibliografía

(1). OMS. (2017). Enfermedades Cardiovasculares. Recuperado de: [https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-disea...](https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-diseases-)(cvds)

(2).Blender, J. (2019). Las enfermedades cerebrovasculares como problema de salud . Revista Cubana de Neurología y Neurocirugía, (2):e335.

(3). Ministerio de Salud Pública-Cuba. (2007). Enfermedades cerebrovasculares, el manejo en la fase aguda y la prevención secundaria del ictus.

MayoClinic. (2023). <https://www.mayoclinic.org/es/diseases-conditions/stroke/symptoms-causes/syc-20350113#:~:text=Un%20accidente%20cerebrovascular%20isquémico%20se,acumulación%20de%20placas%2C%20denominada%20ateroesclerosis>.