

Aplicaciones de la Analítica - Salud

Alanis Álvarez, Juan E. Cardona, Juan E. Soto, Santiago Restrepo

Estudiantes de Ingeniería Industrial

Universidad de Antioquia, Colombia

1. Problemática y diseño de la solución

Problemática de negocio

Un hospital pequeño no cuenta con los recursos para tener un amplio equipo de especialistas en neurología, además, el único especialista disponible con el que cuenta no dispone del tiempo suficiente para realizar la lectura de todas las radiografías de los pacientes de manera oportuna, lo que genera retrasos en los diagnósticos. Esta situación no solo aumenta la insatisfacción de los pacientes, sino que pone en riesgo su salud al posponer el tratamiento adecuado.

Además, este problema está asociado con los costos dado que un diagnóstico a tiempo puede reducir significativamente los costos de tratamiento. En caso de que la detección se retarde, el tumor puede evolucionar y el tratamiento puede generar mayores costos para el hospital (tiempo empleado por el especialista, medicamentos o tratamientos).

Diseño de la solución

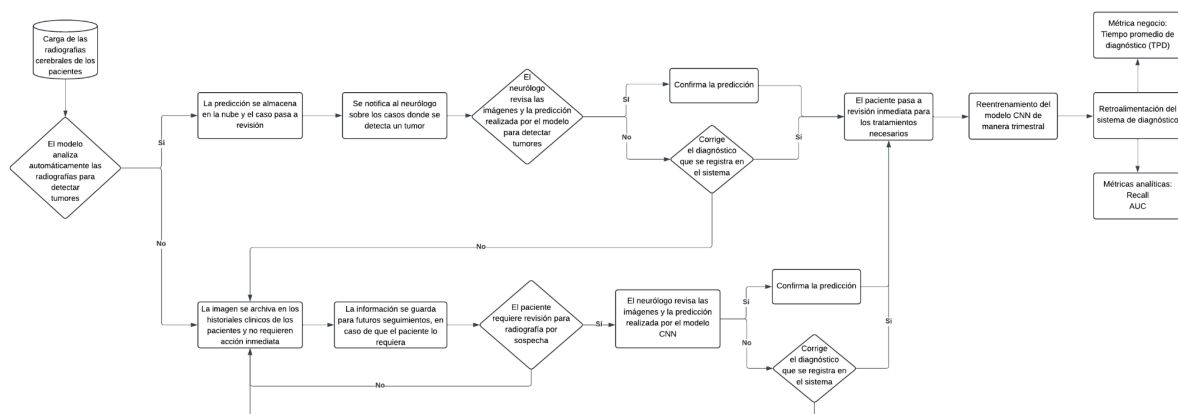


Imagen 1. Diagrama del diseño de la solución caso Salud

Fuente. Elaboración propia en Lucidchart

El montaje de un sistema de diagnóstico de tumores cerebrales asistido por redes neuronales convolucionales (CNN) que automatiza el análisis de imágenes de radiografías cerebrales. Esto funciona como una alternativa que reduce la carga de trabajo del especialista al enfocarse solo en los casos en los que se detecta un posible tumor por el sistema, el tiempo de detección y en últimas, optimiza o agiliza el proceso de diagnóstico de tumores cerebrales. Una vez el sistema (modelo) detecta o predice el tumor y la predicción es positiva, se le notifica al neurólogo para que revise el caso y confirme o corrija la predicción del modelo. Si se confirma la predicción, el paciente pasa a tratamiento inmediato y en caso de corrección, se registra la nueva información en el sistema y se retroalimenta el modelo para mejorar su desempeño, buscando a su vez minimizar la tasa de falsos negativos (el sistema no diagnostica tumor cuando realmente lo tiene).

Para los casos que no se detecta tumor, las imágenes se archivan automáticamente en los historiales clínicos de los pacientes y no requieren una acción inmediata, aunque se pueden utilizar para futuros seguimientos dado el caso que el paciente lo requiera o si le persisten los síntomas. Por otro lado, como se mencionó anteriormente, cuando se detecta un tumor el caso pasa a ser tratado por el especialista.

Adicionalmente, se implementa la integración de los resultados del modelo contenidos en un archivo de Excel con una plataforma en la nube (como Drive) ofreciendo flexibilidad para que el neurólogo realice su trabajo de forma remota, con el propósito de que estos resultados lleguen directamente al especialista sin la necesidad de la disponibilidad física de este en el hospital.

Finalmente el modelo será reentrenado de forma trimestral, teniendo en cuenta la frecuencia de la tasa de correcciones y la cantidad de datos nuevos que surjan.

2. Lectura y preprocesamiento

El conjunto de datos está compuesto por 5712 imágenes para Train y 1311 para Test, divididas en su respectiva clasificación: tumor (clase 1) y no tumor (clase 0). Se realiza una exploración general para visualizar las características de las imágenes con el fin de conocer su tamaño y el valor de la intensidad de los píxeles, donde se realiza el escalado de las imágenes dividiendo su intensidad entre 255. Ahora bien, para conservar la calidad de las imágenes, se realiza una función que busca la imagen con menor resolución y las reescala a (150×198 px), sin embargo, se evidencia que la imagen con menor resolución cuenta con la misma resolución que las demás, es decir, las imágenes tienen buena resolución (ver Imagen 2) y se decide no reescalar.

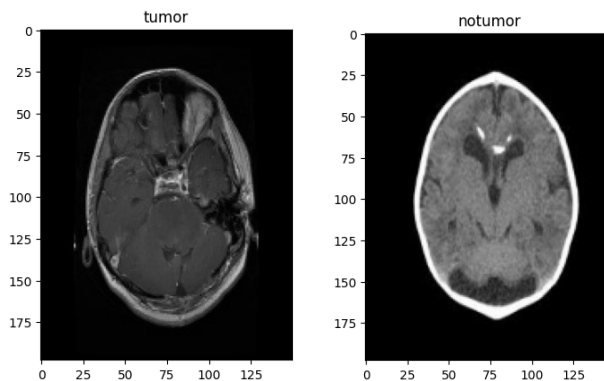


Imagen 2. Vista previa de imágenes en estudio con clasificación tumor y no tumor

3. Selección de variables

Teniendo en cuenta que se realizará la predicción con modelos tradicionales (*machine learning*) y redes neuronales (*deep learning*), es importante mencionar que los tradicionales no admiten el procesamiento de imágenes en 3D, por ende, es necesario realizar la redimensión o aplanamiento de las imágenes, y, para ello se multiplican los píxeles de la imagen de menor dimensión (150×198 px) por sus tres canales RGB, obteniendo como resultado *arrays* de *numpy* en una dimensión con 89.100 elementos para cada imagen. Por otro lado, para el procesamiento de imágenes, por parte de las redes neuronales se mantiene la estructura en tres dimensiones. Finalmente, es importante mencionar que para el montaje de redes neuronales convolucionales es necesario aplicar una reducción de la resolución de

las imágenes debido a dificultades de procesamiento, obteniendo imágenes con una resolución (75×99 px).

4. Métricas

- **Métrica de negocio.** Para la evaluación del impacto de la implementación se hace uso de la siguiente métrica:

$$\text{Tiempo promedio diagnóstico (TPD)} = \frac{\sum \text{Tiempos diagnóstico}}{\text{Total diagnósticos realizados}}$$

Se tiene el objetivo de reducir el valor de esta métrica, dado que en la actualidad, un diagnóstico atraviesa el proceso de espera para ser revisado por parte del especialista, acompañado de la respectiva revisión por parte del mismo. Con la implementación, se espera que el tiempo de espera y diagnóstico se reduzca notablemente, lo que generaría que el tiempo promedio destinado para este proceso sea menor.

- **Métrica analítica.** Dentro de la detección de tumores cerebrales, particularmente se busca reducir al máximo los falsos negativos, dado que pondría en riesgo la vida del paciente al predecir que no presenta tumor sabiendo que sí lo tiene. Para evaluar esta condición se hace uso del *recall* y el *AUC*, donde se espera que estos valores sean lo más cercano posible a 1:

$$\text{Recall} = \frac{\text{Verdaderos positivos}}{\text{Verdaderos positivos} + \text{Falsos negativos}} \quad | \quad \text{AUC: Relación entre verdaderos positivos y falsos negativos}$$

5. Modelos

Inicialmente, se realiza la implementación de los modelos tradicionales, como Regresión Logística, *Random Forest* y *KNN*. Luego se desarrollan dos redes neuronales tradicionales (una con *dropout* y regularización y otra sin *dropout* ni regularización), sin embargo, el objetivo está centrado en hacer uso de redes neuronales convolucionales dado que se cuenta con datos tipo imagen, enfocándonos, como se mencionó anteriormente, en la métrica del *recall* y *AUC*, con el fin de disminuir al máximo la tasa de falsos negativos.

Así mismo, es importante mencionar que al ejecutar y predecir con los modelos tradicionales y con la red neuronal tradicional se obtuvieron métricas óptimas en entrenamiento y un comportamiento parecido en las métricas de prueba que se puede deber a la calidad o alta resolución de las imágenes. Sin embargo, estos modelos no serán tenidos en cuenta dado que el interés se centró en la red neuronal convolucional, la cual requirió de realizar un ajuste de hiperparámetros a través de una grilla de búsqueda para mejorar su desempeño.

6. Optimización de hiperparámetros

La optimización de hiperparámetros se realiza para la red neuronal convolucional buscando la maximización del valor AUC. Este afinamiento permite pasar de un macro average en recall de 86% a 98% en el conjunto de datos de test, y así mismo, en test el AUC pasa de 95% a 99% (ver Tabla 1), lo que permite concluir que el afinamiento de hiperparámetros mejora notablemente el rendimiento del modelo y permite obtener métricas de desempeño de interés superiores. Además, la matriz de confusión (ver Imagen 3) de este modelo deja en evidencia que la tasa de falsos negativos y falsos positivos es realmente baja, lo que permite concluir que el modelo es confiable.

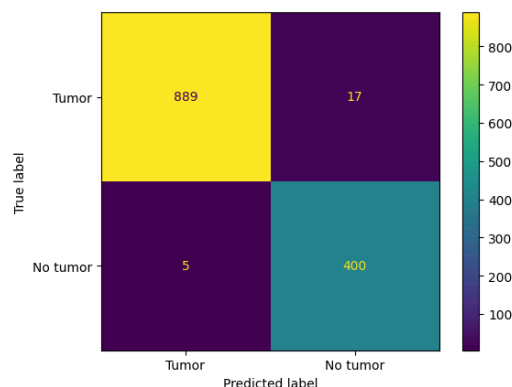


Imagen 3. Matriz de confusión en test para la red neuronal convolucional con afinamiento de hiperparámetros

A continuación, la tabla 1 muestra el desempeño de la red neuronal convolucional en términos de las métricas de *AUC* y *recall*, tanto para cada clase como en general para el conjunto de datos de test, comparando el rendimiento del modelo base con el modelo al cual se le realizó la optimización de hiperparámetros.

Métrica	Sin Tuning (Test)	Con Tuning (Test)
Macro average Recall	0.86	0.98
Recall (clase 0)	0.80	0.99
Recall (clase 1)	0.93	0.98
AUC	0.9456	0.9977

Tabla 1. Comparación de métricas de Recall y AUC en red neuronal convolucional

Los resultados obtenidos después del tuning muestran una mejora significativa en el desempeño de la red neuronal convolucional evidenciada por el aumento en el recall de ambas clases y el incremento del AUC. Gracias a lo anterior, esta red neuronal convolucional optimizada es la elegida para realizar el análisis y el despliegue posterior.

7. Análisis de modelo

Haciendo uso de la red neuronal convolucional con parámetros optimizados mencionada anteriormente se obtiene el siguiente análisis:

- Para clasificados con tumor: Se decide utilizar un threshold de 0.6 dado que al observar la distribución de las probabilidades de clasificación del modelo se observa que este se inclina mayoritariamente por predecir probabilidades más cercanas a 1 que a 0. Así mismo, este threshold permite mantener una tasa de falsos positivos realmente inferior, lo que permite no incurrir en un esfuerzo superior e innecesario por parte del especialista y en tratamientos que no se deben asignar a dichos pacientes. Finalmente, se obtiene un recall en test para la clase 1 (pacientes con tumor) de 98%, lo que se traduce en que el modelo es capaz de predecir 98 de cada 100 pacientes que tienen tumor de manera correcta en el 99% de los casos (1% de error).

- Para clasificados con no tumor: Se decide utilizar un threshold de 0.4, el cual permite acertar el 99% de las predicciones de pacientes que no tienen tumor con un 1% de error. Así mismo, permite mantener una tasa de falsos negativos realmente baja, generando confianza en que el modelo va a predecir correctamente a las personas con tumor, y así, la vida del paciente no correrá riesgo considerable asociado a una no predicción de su tumor.

Finalmente, el modelo permitirá automatizar el 100% de las detecciones de tumores cerebrales del hospital generando una clasificación binaria, donde los pacientes con una probabilidad menor a 0.6 serán clasificados como no tumor y los pacientes con una probabilidad mayor a 0.6 serán clasificados como tumor, de este modo, el modelo será calibrado para no inclinar sus predicciones hacia la clase mayoritaria (clase 1: tumor). Dicho esto, al realizar la predicción con las 1311 imágenes de test utilizando los threshold mencionados anteriormente, se obtiene que el 67.73% (888 pacientes) serán clasificados como tumor y el 32.27% (423 pacientes) restante serán clasificados como no tumor.

8. Despliegue

En cuanto a la automatización del sistema de detección de tumores se hace uso de un conjunto de 337 imágenes de radiografías cerebrales no vistas previamente por el modelo (es decir, no se usaron para entrenamiento ni para test). Para ello, se utiliza el mejor modelo obtenido, la red neuronal convolucional optimizada mediante afinamiento de hiperparámetros, y al mismo tiempo, por limitaciones de capacidad, las imágenes se reducen en resolución, luego se escalan, y se carga el modelo para ejecutar la clasificación binaria. Finalmente, se obtiene como resultado un archivo de Excel (ver vista previa de 7 pacientes en Tabla 2) que incluye la identificación del paciente (el título de la imagen) y la clasificación resultante para cada radiografía, indicando si presenta “tumor” o “no tumor”.

Paciente	Clase
Tr-no 1006	No tumor
Tr-no 1007	No tumor
Tr-piTr 0000	Tumor
Tr-piTr 0001	Tumor
Tr-piTr 0002	Tumor
Tr-piTr 0003	Tumor
Tr-piTr 1006	Tumor

Tabla 2. Predicciones generadas en Excel por el modelo para datos nuevos en despliegue

Para concluir, este archivo de excel generado en el despliegue del sistema será el que se integre con una plataforma como Google Drive para que el especialista pueda tener los resultados del modelo al alcance de su mano en cualquier momento que lo requiera.

9. Conclusión

Una vez implementado el proceso, se puede concluir que el hospital se vería enormemente favorecido con la puesta en marcha del sistema de detección de tumores cerebrales de forma automática haciendo uso de una red neuronal convolucional. Esto, dado que, tanto la carga de trabajo del especialista como el tiempo promedio de diagnóstico se reducen, y así mismo, se aumenta la posibilidad de atención de pacientes en tiempo real.

Por otro lado, debido a que el sistema se encuentra en una fase de implementación se considera fundamental que el neurólogo revise y confirme las predicciones clasificadas como tumor. Además, a pesar de que el modelo utilizado cuenta con una tasa baja de falsos negativos (pacientes con tumor que el modelo predice como no tumor), si el paciente requiere de una nueva revisión de su radiografía previa, debido a sospechas en su diagnóstico, el especialista es el encargado de ejecutarla y confirmar la predicción realizada, lo que genera un alto grado de confiabilidad en los resultados del sistema.

Enlace del repositorio: <https://github.com/SantiRestrepoOlarte/ProyectoHealth.git>