



º



Detección automatizada deL glioblastoma multiforme

Descargo parcial PBL



29 de Enero de 2021

Ainhoa arruabarrena ortiz

Edgar azpiazu crespo

ÍNDICE

[1. INTRODUCCIÓN 1](#_Toc59085392)

[2. ESTADO DEL ARTE 3](#_Toc59085393)

[2.1. Glioblastoma multiforme 3](#_Toc59085394)

[2.2. Machine Learning en medicina 4](#_Toc59085395)

[2.2.1. Aprendizaje supervisado 5](#_Toc59085396)

[2.2.2. Aprendizaje no supervisado 7](#_Toc59085397)

[2.2.3. Evaluación de los sistemas 8](#_Toc59085398)

[2.2.4. Aplicaciones 11](#_Toc59085399)

[2.3. Machine Learning Glioblastoma 13](#_Toc59085400)

[2.3.1. Técnicas tradicionales de Machine Learning 13](#_Toc59085401)

[2.3.2. Nuevas técnicas de Deep Learning 14](#_Toc59085402)

[3. OBJETIVOS 16](#_Toc59085403)

[4. FASES 17](#_Toc59085404)

[4.1. Planificación 18](#_Toc59085405)

[5. DISEÑO DE LA SOLUCIÓN 20](#_Toc59085406)

[5.1. Requisitos técnicos 20](#_Toc59085407)

[5.2. Herramientas 21](#_Toc59085408)

[6. DESARROLLO 22](#_Toc59085409)

[7. BIBLIOGRAFÍA 23](#_Toc59085410)

ÍNDICE DE FIGURAS

[Figura 1. Proceso del diagnóstico [1]. 1](#_Toc59085411)

[Figura 2. Ejemplo de aprendizaje no supervisado [15]. 7](#_Toc59085412)

[Figura 3. Matriz de confusión [16]. 9](#_Toc59085413)

[Figura 4. Curva PR para distintos valores de umbral [19]. 10](#_Toc59085414)

[Figura 5. Ejemplo de curva ROC [19]. 11](#_Toc59085415)

[Figura 6. Histogramas de la distribución de las características más predictivas de la esperanza de vida [31]. 14](#_Toc59085416)

[Figura 7. Flujo de trabajo del análisis radiómico [33]. 15](#_Toc59085417)

[Figura 8. Diagrama de Gantt del proyecto. 19](file:////Users/ainhoaarruabarrenaortiz/Documents/Master/PBL/pbl_glioblastoma/Doc/Descargo%20Parcial/Descargo%20Parcial%20PBL%20-%20Ainhoa%20-%20Edgar.docx#_Toc59085418)

ÍNDICE DE TABLAS

[Tabla 1. Resultados de los algoritmos de aprendizaje supervisado [16]. 6](#_Toc59085419)

# INTRODUCCIÓN

El proceso de diagnóstico es una actividad compleja y colaborativa que pide de un razonamiento clínico y recogida de información con el objetivo de determinar el estado de salud del paciente. Es un proceso (Figura 1) que se extiende en el tiempo, siendo iterativo y mientras se da la recogida de información, el objetivo es reducir las dudas sobre el diagnóstico, limitar las distintas posibilidades para el resultado del diagnóstico y desarrollar un entendimiento más preciso y completo del problema de salud del paciente.

Diagram

Description automatically generated

Figura 1. Proceso del diagnóstico [1].

La obtención de un diagnóstico correcto es clave, ya que proporciona información sobre el estado del paciente y es la información sobre la cual se basan los tratamientos posteriores para la mejora de este. Los servicios de salud cuentan con un error común para todos los ámbitos de estudio, los errores en el diagnóstico, que siguen produciéndose en una alta proporción.

Es probable que la mayoría de las personas experimente al menos un error de diagnóstico en su vida, a veces con consecuencias devastadoras. Los errores de diagnóstico pueden causar daño a los pacientes al impedir o retrasar el tratamiento adecuado, proporcionar un tratamiento innecesario o perjudicial, o tener repercusiones psicológicas o financieras. Por ejemplo, se han encontrado los siguientes datos [1].

* Un cálculo conservador encontró que el 5% de los adultos de EE. UU. que buscan atención médica cada año experimentan un error de diagnóstico [2].
* La investigación durante décadas sobre las autopsias ha demostrado que los errores de diagnóstico contribuyen a aproximadamente el 10% de las muertes de pacientes.
* Las revisiones de los registros médicos sugieren que los errores de diagnóstico representan entre el 6 y el 17 % de los eventos adversos del hospital.
* Los errores de diagnóstico son el principal tipo de reclamaciones por negligencia médica pagadas, tienen casi el doble de probabilidades de haber resultado en la muerte del paciente en comparación con otras reclamaciones, y representan la mayor proporción de los pagos totales.

A pesar de la presencia generalizada de los errores de diagnóstico y del riesgo de que se produzcan graves daños a los pacientes, estos no han sido muy considerados a la hora de mejorar la calidad y seguridad de los pacientes en los servicios de salud. Sin un enfoque dedicado a mejorar el diagnóstico, estos errores probablemente empeorarán a medida que la prestación de la atención de la salud y el proceso de diagnóstico sigan aumentando en complejidad.

Teniendo todo esto en cuenta, se quiere desarrollar un sistema basado en inteligencia artificial (IA) que sea capaz de detectar un Glioblastoma partiendo de imágenes médicas adquiridas mediante RM o TAC. Asimismo, se quieren combinar las características del tumor extraídas de las imágenes como pueden ser su posición, tamaño, forma o textura con otros datos poblacionales y hábitos de vida sobre los pacientes para tratar de encontrar alguna relación con la aparición del Glioblastoma.

# ESTADO DEL ARTE

En esta sección se expone información sobre la enfermedad bajo estudio, el Glioblastoma multiforme, así como un resumen de Machine Learning (ML) y sus aplicaciones en el ámbito médico, información necesaria para la comprensión de la motivación de este proyecto.

## Glioblastoma multiforme

El cáncer es el nombre que se da a un conjunto de enfermedades relacionadas. En todos los tipos de cáncer, algunas de las células del cuerpo empiezan a dividirse sin detenerse y se diseminan a los tejidos del derredor [3].

Dentro de los distintos tipos de cáncer, Glioma es el término utilizado para referirse a los tumores cerebrales primarios que, a su vez, se clasifican dependiendo de la célula que lo origina. Son los tumores más comunes del sistema nervioso central (SNC) y agrupan casi el 80% de todos los tipos de tumores del cerebro. En este grupo se incluye el Glioblastoma multiforme, el tumor más maligno [4]. Existen cuatro grados histológicos del Glioma. Los Glioblastomas pertenecen al grado 4 y son los más comunes (53%). Los Gliomas difusos de grado 2 (difusos de bajo grado) y grado 3 (anaplásticos) representan aproximadamente el 30% de todos los Gliomas. La edad media en el momento del diagnóstico de estos Gliomas es de 64, 43 y 56 años, respectivamente. Por otra parte, los Gliomas pediátricos más comunes son los de grado 1 (predominantemente astrocitomas pilocíticos) que representan el 33% de los Gliomas pediátricos [5].

Los principales tipos de tratamiento actualmente consisten en una combinación entre una cirugía que tiene como objetivo la resección máxima del tumor, radioterapia y quimioterapia. A pesar de estos tratamientos y los avances realizados, siguen sin ser muy eficaces y los pacientes tienen una esperanza de vida alrededor de 16-18 meses y tan solo el 3-5% sobrevive los 5 años [6]. A pesar de que su incidencia sea de menos de 10 por 100000 personas, su mal pronóstico lo convierte en un problema crucial de salud pública. Puede darse a cualquier edad, pero tiene mayor incidencia en el rango de 55 a 60 años.

La detección del Glioblastoma se basa tanto en técnicas invasivas como técnicas de imagen no invasivas. El método por excelencia [7] para detectar los tumores son los escáneres de resonancia magnética (RM) debido al contraste superior que ofrecen en tejidos blandos, lo que permite visualizar mejor la complejidad y heterogeneidad del tumor. En el caso de que el paciente no pueda someterse a una resonancia magnética, se suele recomendar el uso de una tomografía computarizada (TAC).

Una detección precoz del cáncer es de suma importancia. La detección y diagnóstico mediante técnicas de imagen avanzadas reduce la necesidad de una cirugía exploratoria para ello. Además, la posibilidad de detectar y tratar un glioma de bajo grado, antes de que progrese a un Glioblastoma, puede marcar la diferencia entre la vida y la muerte [8]. Así, se podrían tratar cuando todavía son pequeños antes de que se conviertan en malignos y aún pueden ser curables mediante cirugía mínimamente invasiva o incluso por métodos de radiación estereotáctica como la radiocirugía [9]. Cabe destacar que no siempre es posible detectar un glioma de bajo grado ya que a veces se generan de nuevo sin evidencias de un precursor de menor grado [10]. La fundación del Glioblastoma afirma que una detección precoz podría llegar a posibilitar un tratamiento sin cirugía. También expone que varios estudios han demostrado que la resección del Glioblastoma puede estimular cualquier célula cancerígena no eliminada y hacer que crezca hasta un 75% más rápido que antes de la intervención, dando aún más importancia a la detección [11].

Los datos que se extraen para el diagnóstico como para el seguimiento de esta enfermedad se realizan de forma manual habitualmente, siendo evaluados por profesionales para la toma de decisiones. En un futuro, el objetivo es que la inteligencia artificial (IA) sea la encargada de realizar un primer diagnóstico de una forma más precisa a la que pueda realizar el profesional, sirviendo de base para un mejor diagnóstico y formando parte de la experiencia de la medicina personalizada. Hoy en día existen sistemas de este tipo aplicadas a la oncología, como por ejemplo el sistema IBM Watson, encargado de asistir a los médicos a la hora de planificar el tratamiento mediante el análisis de datos clínicos, datos genéticos e imágenes médicas, realizando recomendaciones [12].

## Machine Learning en medicina

El aumento de conocimiento y comprensión de las enfermedades ha conllevado el crecimiento de la información y datos asociados a estos, en parte gracias a los avances que ha habido en herramientas de análisis cuantitativo y cualitativo en cuanto a mediciones de parámetros fisiológicos. Al contar con un campo de datos tan grande, este se convierte en un escenario ideal para el uso de técnicas de Machine Learning (ML). El potencial de herramientas de ML como plataforma para la obtención de información de numerosas fuentes en un sistema para la asistencia en la toma de decisiones está siendo descubierto [13].

Todo proceso de aprendizaje consiste en dos fases distintas que se definen a continuación [14].

1. Estimación de las dependencias desconocidas de un sistema a partir de un conjunto de datos determinado
2. Empleo de las dependencias estimadas para predecir los nuevos productos del sistema.

El objetivo principal de las técnicas de ML es producir un modelo que pueda utilizarse para realizar una clasificación, predicción, estimación o cualquier otra tarea similar. Los algoritmos de ML existentes se pueden clasificar en dos grupos: aprendizaje supervisado y aprendizaje no supervisado. Además, existe la posibilidad de aplicar algoritmos de aprendizaje semi-supervisado, lo que combina elementos de los dos grupos.

### Aprendizaje supervisado

El aprendizaje supervisado se refiere a técnicas donde un modelo se entrena con una serie de entradas o características asociadas a un resultado relevante, con el objetivo de que cuando el modelo esté entrenado, este sea capaz de predecir el resultado mediante la introducción de nuevos datos. Los resultados de las predicciones de los modelos del aprendizaje supervisado pueden ser discretos o continuos.

A veces, para que el algoritmo de ML sea capaz de comprender la información de las entradas, por ejemplo, en el reconocimiento de imágenes o procesamiento del lenguaje, las variables (píxeles o palabras en este caso) deben ser procesadas para extraer de estos una serie de características identificables que puedan ser representadas en una matriz numérica. En el caso de las imágenes, una característica puede ser el color del píxel y en el caso de las palabras, el número de veces que aparece en un texto.

Una vez que un conjunto de datos se ha organizado en características y resultados, se le puede aplicar un algoritmo de ML. El algoritmo se mejora iterativamente para reducir el error de predicción mediante una técnica de optimización. [15]

Es importante tener en cuenta que al entrenar los algoritmos de ML es posible sobreajustar el algoritmo a un conjunto de datos específico, lo que da lugar a un modelo de predicción que no generaliza bien a los nuevos datos. Esto se denomina *overfitting* y el riesgo de que este se de puede reducirse mediante el uso de bases de datos de mayor tamaño, reducción del número de características seleccionando solo las más importantes o gracias a las técnicas de división de los datos que se explican en el apartado 2.2.3 Evaluación de los sistemas.

Se distinguen dos tipos de algoritmos dentro del aprendizaje supervisado.

#### Clasificación

Un modelo que produce categorías discretas (a veces denominadas clases) se denomina algoritmo de clasificación. Por ejemplo, mediante este tipo de algoritmos se puede determinar si un tumor es benigno o maligno.

En la práctica, los algoritmos de clasificación devuelven la probabilidad de una clase (entre 0 para imposible y 1 para definitiva). Típicamente, transformaríamos cualquier probabilidad superior a 0,50 en una clase de 1, pero este umbral puede ser alterado para mejorar el rendimiento del algoritmo según sea necesario.

Existen distintos tipos de algoritmos que entran dentro de este grupo, algunos se nombran a continuación [16] y sus rendimientos se comparan en la Tabla 1.

* *Decision Trees*: se emplean para organizar gráficamente la información sobre posibles opciones, consecuencias y valor final. Se utilizan para decidir entre varios cursos de acción. Crean una representación visual de los diversos riesgos, recompensas y valores potenciales de cada opción.
* Clasificación mediante Naive Bayes: método probabilístico para la clasificación multiplicando las probabilidades individuales de cada par de atributos-valor. Este simple algoritmo asume la independencia entre los atributos.
* Redes neuronales artificiales (Artificial Neural Networks o ANN): sistema diseñado para similar la forma en la que el ser humano analiza y procesa la información. Se construyen mediante capas de neuronas interconectadas entre ellas. Normalmente una única capa es suficiente para la clasificación de la mayoría de conjunto de datos.
* *Support Vector Machines* (SVN): es una función que toma el espacio de entrada de baja dimensión y lo transforma en un espacio de mayor dimensión, es decir, convierte un problema no separable en un problema separable. Es muy útil cuando se trata de problemas donde la separación a realizar no es lineal y se puede emplear también para problemas de regresión.

Tabla 1. Resultados de los algoritmos de aprendizaje supervisado [16].

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | *Decision Trees* | Naive Bayes | Redes Neuronales | SVN |
| *Accuracy* | \*\* | \* | \*\*\* | \*\*\*\* |
| Velocidad aprendizaje | \*\*\* | \*\*\*\*\* | \* | \* |
| Velocidad de clasificación | \*\*\* | \*\*\* | \*\*\*\* | \*\*\*\* |
| Manejo de atributos irrelevantes | \*\*\* | \*\* | \* | \*\*\*\* |
| Tolerancia a atributos redundantes | \*\* | \* | \*\* | \*\*\* |
| Prevención de *overfitting* | \*\* | \*\*\* | \* | \*\* |

#### Regresión

Un modelo que devuelve una predicción de un valor continuo se conoce como algoritmo de regresión. El uso del término regresión en el ML varía de su uso en la estadística, donde la regresión se utiliza a menudo para referirse tanto a los resultados binarios (es decir, la regresión logística) como a la salida continua (es decir, la regresión lineal). Por ejemplo, este tipo de modelos se pueden emplear para predecir la esperanza de vida de un individuo o la dosis tolerable de quimioterapia.

### Aprendizaje no supervisado

El aprendizaje no supervisado se refiere a técnicas donde el modelo se entrena con una serie de entradas o características que no tienen asociadas un resultado o etiqueta de la clase a la que pertenecen. De esta forma, el objetivo del sistema es desarrollar y organizar los datos, buscando características comunes entre ellos, y cambiando en base a los conocimientos internos.

Estas técnicas se denominan a menudo técnicas de reducción de dimensiones. A veces se emplean métodos no supervisados junto con los métodos de aprendizaje supervisado, ya que, al comprimir la información de un conjunto de datos en menos características o dimensiones, se pueden evitar problemas como el alto costo de los cálculos. [15]

En la Figura 2 se muestra visualmente una técnica de reducción de dimensiones. En esta figura, los datos originales se muestran a la izquierda representados por distintas formas y se introducen en un algoritmo de aprendizaje no supervisado que agrupa los datos. Las características que no tienen suficiente relevancia se excluyen, reduciendo así el número de características dentro del conjunto de datos y permitiendo la agrupación.

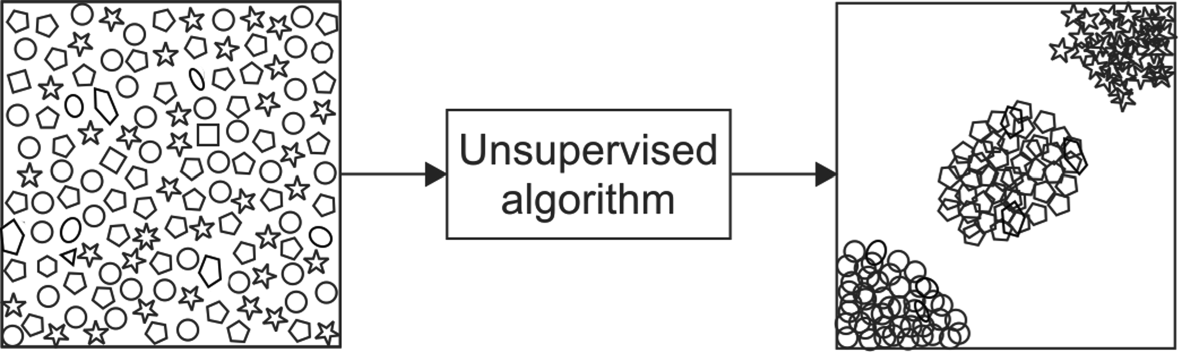


Figura 2. Ejemplo de aprendizaje no supervisado [15].

Los algoritmos principales de se presentan a continuación.

* *K-means clustering*: comienza con un primer grupo de centroides seleccionados al azar, que se utilizan como puntos de inicio para cada clúster, y luego realiza cálculos iterativos para optimizar las posiciones de los centroides, deteniéndose solo cuando los centroides se han estabilizado (agrupación exitosa) o se ha llegado al número máximo de iteraciones.
* PCA (*Principal Component Analysis*): técnica utilizada para reducir la dimensionalidad de los conjuntos de datos de alta dimensión, conservando la estructura original y las relaciones inherentes al conjunto de datos original.
* ICA (*Independent Component Analysis*): técnica para la obtención de bases linealmente independientes para los datos, es decir, identifica componentes o fuentes independientes de los datos.
* ANN: aprende a categorizar los datos en base la distancia entre los clústeres dentro de ella. La técnica empleada se denominan mapas autoorganizados y se trata de mapas computacionales topográficos de inspiración biológica que aprenden mediante la auto-organización de sus neuronas [17].

Estos algoritmos de aprendizaje no supervisado pueden dividirse en dos grupos.

#### *Clustering*

Un problema de *clustering* se presenta cuando se quiere descubrir las agrupaciones inherentes de los datos.

#### Asociación

Un problema de aprendizaje de reglas de asociación es cuando se quiere descubrir reglas que describan grandes porciones de los datos. Por ejemplo, una regla de asociación puede ser que las personas que compran X también tienden a comprar Y.

### Evaluación de los sistemas

Los datos empleados para la fase de entrenamiento del modelo no pueden ser empleados para la evaluación de este, dado que el sistema recordará esa combinación concreta y realizará una predicción correcta para todos los datos. Para poder realizar la evaluación de los sistemas obtenidos de forma correcta, existen dos métodos de separación de datos distintos para la creación del modelo [18].

* *Holdout*: divide los datos en 3 grupos distintos con distintos porcentajes. Un conjunto de datos se emplea para el entrenamiento del modelo y su creación, un segundo conjunto para la validación, y un tercero para el testeo del modelo tras su creación. Esta técnica se emplea dado su simplicidad, velocidad y flexibilidad, pero se asocia a una alta variabilidad, ya que las diferencias entre el conjunto de datos empleados para el entrenamiento y el testeo del modelo pueden afectar a la precisión de las predicciones.
* *Cross-validation*: divide los datos en un grupo para el entrenamiento y otro para la evaluación. La técnica más habitual se denomina *k-fold*, y consiste en dividir el conjunto de datos original en k grupos distintos del mismo tamaño, a los cuales se les denomina *folds*. El entrenamiento del modelo se repite *k* veces, empleando cada vez un conjunto para la validación y el resto para el entrenamiento. La estimación del error se promedia sobre todas las combinaciones para obtener la eficacia total de nuestro modelo.

Para poder evaluar correctamente el rendimiento del sistema se definen distintos parámetros: True Positive (TP), True Negative (TN), False Positive (FP) y False Negative (FN) que se expresan de forma gráfica en la Figura 3, lo que se denomina matriz de confusión. Mediante estos parámetros se obtienen las distintas métricas que se explican a continuación.

Graphical user interface, diagram

Description automatically generated

Figura 3. Matriz de confusión [16].

#### Accuracy (ACC)

Muestra el porcentaje de predicciones realizadas correctamente por el sistema. Se obtiene mediante la fórmula definida en la ecuación (1).

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

#### Sensitivity o Recall (True Positive Rate o TPR)

Se define como la capacidad de clasificar los elementos positivos como tal y se obtiene mediante la ecuación (2).

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

#### Specificity (True Negative Rate o TNR)

Se define como la capacidad de clasificar los elementos negativos como tal y se obtiene mediante la ecuación (3).

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

#### *Precision* o valor predictivo positivo (*Positive Predictive Value* o PPV)

Expresa cuales de los casos clasificados como positivos, cuales se corresponden en la realidad con un caso positivo. Se obtiene mediante la ecuación (4).

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

#### Valor predictivo negativo (*Negative Predictive Value* o NPV)

Expresa cuales de los casos clasificados como negativos, cuales se corresponden en la realidad con un caso negativo. Se obtiene mediante la ecuación (5).

|  |  |  |
| --- | --- | --- |
|  |  | (5) |

#### Valor F1

Se trata de la media harmónica entre el valor de *precision* y *recall*. De esta forma con una única métrica se evalúan ambos parámetros y de esta forma, cuanto mayor sea el valor de F1, mejor. Se calcula de la forma expresada en la ecuación (6).

|  |  |  |
| --- | --- | --- |
|  |  | (6) |

Uno de los problemas que presenta esta métrica es que *precision* y *recall* tienen el mismo peso, por lo que depende de la aplicación habrá que buscar otras métricas para mejorar esta interpretación.

#### Curva PR

Se trata de la curva entre la *precision* y el *recall* para distintos umbrales de probabilidad, lo que permite el estudio del equilibrio entre la tasa de positivos reales y la tasa de negativos reales. Un ejemplo de esta curva se observa en la Figura 4. De esta forma, la parte superior izquierda de la gráfica es el espacio ideal donde ambos valores son los más altos posibles. A partir esta curva, también se calcula el *Area Under the Curve* (AUC), que sirve como resumen del rendimiento del modelo, siendo este mejor cuanto mayor sea el valor del AUC.

Chart

Description automatically generated

Figura 4. Curva PR para distintos valores de umbral [19].

#### Curva ROC

Herramienta útil para la evaluación de sistemas de clasificación binarios. La curva ROC (*Receiver Operating Characteristic*) y representa el TPR o *recall* y el FPR (*False Positive Rate*), cuya formula está representada por la ecuación que se encuentra en (7).

|  |  |  |
| --- | --- | --- |
|  |  | (7) |

En la curva ROC, la parte superior izquierda de la gráfica representa una clasificación perfecta realizada por el sistema (Figura 5). De esta forma se compara el rendimiento del sistema con los resultados de un sistema donde la clasificación se realiza de forma aleatoria (representada mediante una línea roja en la Figura 5). A partir esta curva también se calcula el AUC, que sirve como resumen del rendimiento del modelo, siendo este mejor cuanto mayor sea el valor del AUC.

Chart, line chart

Description automatically generated

Figura 5. Ejemplo de curva ROC [19].

### Aplicaciones

Dos de las áreas donde el uso de ML puede ser beneficioso dentro del campo de la medicina son el diagnostico y predicción del resultado. Esto incluye la posibilidad de identificar algunas emergencias médicas como pueden ser las recaídas o las transiciones a enfermedades distintas.

Algunas de las aplicaciones de ML que se usan hoy en día en un contexto médico se expresan a continuación [20].

#### Chatbot

Programa informático con el cual se puede mantener una conversación gracias a la IA. Aplicados en este contexto, cuentan con capacidad para reconocer el lenguaje e identificar patrones en la expresión de los pacientes para realizar un diagnóstico potencial, prevenir enfermedades o recomendar un curso de acción apropiado.

Por ejemplo, una startup con sede en U.K. ha desarrollado la aplicación *Babylon Health* [21], que mediante el reconocimiento del lenguaje, compara los síntomas que recibe de un usuario con una base de datos de enfermedades. En respuesta, recomendará un curso de acción apropiado basado en una combinación de los síntomas reportados, el historial del paciente y sus circunstancias. La aplicación también está diseñada para integrar los datos de los dispositivos del paciente para controlar los signos vitales como el ritmo cardíaco y el nivel de colesterol.

#### Patología

Se basa en el análisis de laboratorio de los fluidos corporales como la sangre y la orina, así como de los tejidos, para realizar diagnósticos de enfermedades. Se emplea la visión artificial y tecnologías de IA para realizar estas tareas que hasta este momento solo eran capaz de hacer patólogos mediante el uso de microscopios.

Con el objetivo mejorar la velocidad y la precisión de los diagnósticos, un equipo de investigadores procedentes del Centro Médico Beth Israel Deaconess y de la Facultad de Medicina de Harvard han utilizado DL para entrenar un algoritmo capaz de integrar el reconocimiento de lenguaje y de imágenes para diagnosticar tumores en tejido. Este algoritmo demostró un 92% del éxito, algo menos que el 96% de acierto con el que cuentan los profesionales. Sin embargo, combinando el uso del algoritmo y el diagnóstico por parte de profesionales, el porcentaje del acierto obtenido asciende al 99,5% [22].

#### Enfermedades raras

Se emplea software de reconocimiento facial con técnicas de ML como ayuda al diagnóstico de enfermedades raras. Las fotos de los pacientes se analizan mediante el análisis facial y mediante DL se detectan fenotipos que se correlacionan con enfermedades genéticas raras.

La aplicación Face2Gene [23] es un software de reconocimiento facial se está combinando con el aprendizaje de la máquina para ayudar a los médicos a diagnosticar enfermedades raras (en este caso, de los rasgos dismórficos faciales). Las fotos de los pacientes se analizan mediante el análisis facial y mediante DL se detectan fenotipos que se correlacionan con enfermedades genéticas raras. La plataforma está actualmente disponible sólo para profesionales y así prevenir falsos positivos, soportando más de 7.500 trastornos.

#### Oncología

DL se está empleando para entrenar algoritmos para diversas funciones, como segmentar un tumor o realizar un seguimiento del desarrollo de un tumor, a un nivel comparable al de los profesionales médicos.

Investigadores de la Universidad de Stanford han desarrollado un algoritmo empleando DL, en concreto redes neuronales convolucionales profundas (CNN), para el diagnóstico de cáncer de piel. Entrenan el algoritmo con un conjunto de datos de 129.450 imágenes clínicas, que se traducen en 2.032 enfermedades diferentes. Su desempeño se comprobó contra 21 dermatólogos certificados en imágenes clínicas probadas por biopsia, demostrando que el algoritmo desarrollado logra un rendimiento a la par de todos los expertos [24].

## Machine Learning Glioblastoma

En las últimas décadas se ha realizado una evolución continua en lo que respecta a la investigación sobre el cáncer. Se han aplicado distintos métodos, por ejemplo, la detección temprana, para identificar los tumores antes de que causen síntomas, creando a su vez nuevas estrategias.

Con la introducción de las nuevas tecnologías en la medicina, se han recopilado grandes cantidades de datos sobre cáncer a disposición de la investigación científica. Sin embargo, la predicción precisa de los resultados de una enfermedad sigue siendo un reto para los médicos, lo que ha provocado que métodos de Machine Learning (ML) se hayan convertido en una herramienta popular para investigadores en este campo de estudio. Estas técnicas pueden descubrir e identificar patrones y relaciones entre ellos a partir de complejos conjuntos de datos, que a simple vista pueden pasar desapercibidos por el análisis de un profesional. Además, algunas técnicas son capaces de predecir eficazmente los resultados futuros de un tipo de cáncer.

La segmentación manual de los volúmenes tumorales requiere mucho tiempo y sufre de una gran variabilidad de un caso a otro. Esto limita su uso en la práctica clínica. Los métodos asistidos por ordenador reducen el tiempo de segmentación y demuestran una buena concordancia con segmentaciones manuales verificadas [25].

En los últimos años se han desarrollado nuevos enfoques de IA para gliomas con los que poder predecir su grado y genómica a partir de imágenes, automatizar su diagnóstico y proporcionar información sobre su pronóstico [26]. Técnicas nuevas de DL, así como redes neuronales y redes neuronales convolucionales (CNN) han facilitado la extracción automatizada de datos clínicos relevantes para la planificación del tratamiento y la monitorización posterior al mismo [27].

### Técnicas tradicionales de Machine Learning

La radiómica, la extracción de alto rendimiento de características de imágenes cuantitativas tiene el potencial de proporcionar información sobre la fisiopatología subyacente y de desarrollar biomarcadores que son difíciles de descubrir mediante la inspección visual. Normalmente, la radiómica incluye las siguientes fases: preprocesamiento de las imágenes, estimación de características, selección de características, clasificación y evaluación [5]. Varios estudios han demostrado que la combinación de la radiómica con técnicas tradicionales de ML es capaz de diferenciar el Glioblastoma de otros tipos de cáncer como un linfoma o tumores cerebrales metastásicos [28], [29], prediciendo información genética sobre el tumor, así como su posible respuesta ante el tratamiento o la posibilidad de sobrevivir a él [30].

Se han hecho estudios basados en características extraídas de imágenes de RM con las que, mediante un algoritmo de ML, poder predecir la expectativa de vida y el subtipo molecular del tumor. Utilizando la validación cruzada se asegura la generalización de los pronosticadores creados a nuevos pacientes. A pesar de las diversas correlaciones entre las características de las imágenes, la expresión genómica y la supervivencia reportadas en la literatura, ningún análisis ha sido lo suficientemente contundente como para entrar en la práctica clínica. Por lo tanto, existe una dificultad considerable a la hora de determinar los biomarcadores que han de usarse y cuáles proporcionarán el resultado más exacto y reproducible. Por ejemplo, la Figura 6 muestra la distribución de las características más predictivas que marcan la supervivencia del paciente.

Diagram

Description automatically generated

Figura 6. Histogramas de la distribución de las características más predictivas de la esperanza de vida [31].

Como demuestran los histogramas, es frecuente que existan pequeñas diferencias entre los supervivientes a corto y a largo plazo, y sólo surgen patrones distintivos que predicen la supervivencia mediante la integración de múltiples rasgos. Teniendo esto en cuenta, se han logrado predecir de forma no invasiva y precisa las variables clínicas y genómicas de los pacientes. La esperanza de vida de los pacientes (corta, media, larga) se ha predicho con un *accuracy* del 80% mientras que en la clasificación del subtipo molecular se ha obtenido un 76% [31].

También se han realizado estudios para tratar de mejorar la clasificación de los tejidos de Glioblastoma heterogéneos. Así, se ha predicho el grado del Glioma presente en los pacientes. Para ello, se han extraído varias características haciendo uso de radiómica para, después, seleccionar las más significativas. Tras esto, se ha entrenado y probado un método de clasificación usando árboles de decisión *random forest*. Se ha obtenido una predicción consistente del grado del tumor en el 96% de los casos analizados [32].

### Nuevas técnicas de Deep Learning

La extracción de características relevantes es una tarea crítica en la radiómica. Muchos estudios utilizan características extraídas a mano o explícitamente diseñadas. Estas incluyen la forma, intensidad y textura de los tumores entre otras. Aun así, estos rasgos de imagen son poco profundos y de bajo orden y es posible que no caractericen plenamente la heterogeneidad de la imagen, por lo que pueden limitar el potencial del modelo radiómico. El aprendizaje por transferencia puede incorporarse al modelo radiómico actual para la extracción de características profundas de varias capas de una CNN. El flujo del análisis radiómico se expone en la Figura 7.

Diagram

Description automatically generated

Figura 7. Flujo de trabajo del análisis radiómico [33].

Estos rasgos profundos contienen información más abstracta de las imágenes médicas y pueden proporcionar más patrones de predicción en comparación con los rasgos extraídos a mano [33].

Las CNN son el método asistido por ordenador puntero para la segmentación del Glioblastoma. Superan a los métodos alternativos que utilizan clasificadores de decisión independientes para extraer características de textura e intensidad. Se utiliza un conjunto de datos pre-anotados para entrenar la arquitectura de las CNN y realizar una serie de convoluciones matemáticas a través de múltiples capas interdependientes. Así se determina la relación entre las imágenes de entrada y las de salida. Después, la CNN puede ser validada en distintos conjuntos de datos de prueba. Por ejemplo, DeepMedic es una CNN tridimensional con una alta calificación en el desafío de segmentación de tumores cerebrales (BRATS). DeepMedic asigna clases a cada vóxel de forma independiente usando la intensidad y la información de características locales. Ha demostrado que las mediciones volumétricas de las subregiones del tumor pueden medirse con gran precisión a partir de imágenes preoperatorias de RM. Además, la CNN puede integrarse en el flujo de trabajo radiológico para acortar el tiempo de segmentación hasta 20 minutos por paciente comparado con una segmentación completamente manual [25].

# OBJETIVOS

El objetivo principal del proyecto es desarrollar un sistema inteligente que realice las siguientes funciones.

* Obtener un primer diagnóstico mediante la detección del tumor desde las imágenes médicas y obtener una clasificación de este.
* Predecir la probabilidad de supervivencia del paciente.
* Pronosticar las posibilidades de tener una recaída tras la eliminación del tumor.

Todo esto deberá desarrollarse basándose en el análisis de imágenes médicas y parámetros poblacionales extraídos de pacientes que sufren de dicha enfermedad. Se han planteado una serie de objetivos secundarios los cuales definen los elementos necesarios para lograr cumplir el objetivo principal y se exponen a continuación.

* El sistema por desarrollar debe basarse en datos obtenidos de al menos 50 pacientes para asegurar la fiabilidad del sistema de predicción.
* El sistema deberá evaluarse empleando distintos parámetros y pruebas estadísticas. El sistema tiene que ser capaz de realizar las predicciones con una sensibilidad del 0.7, especificidad del 0.7, *accuracy* del 70%, VPP y VPN del 70% y AUC del ROC del 0.8.
* El sistema deberá ser capaz de localizar el tumor con una precisión de ±0.1mm.
* Identificar las variables más relevantes que contribuyen a la predicción.

Como objetivo adicional, si se logran todos los puntos anteriores dentro de los límites marcados, los resultados obtenidos en la predicción y el diagnóstico, así como, en caso de tener un diagnóstico desfavorable, la imagen médica con el Glioblastoma detectado, deben mostrarse al usuario en una interfaz gráfica. Esta interfaz también debe posibilitar la introducción de nuevos pacientes para el diagnóstico. Además, dicha interfaz debe ser intuitiva y fácil de usar para todo tipo de usuarios.

# FASES

El proyecto ha sido dividido en distintas fases que marcarán el desarrollo del proyecto y son necesarias para alcanzar el objetivo fijado.

1. Elaboración de un estado del arte sobre la detección del Glioblastoma multiforme y de las técnicas empleadas para su diagnóstico. El estado del arte deberá contener información sobre los siguientes puntos.
   * Información sobre el impacto que tiene este diagnóstico, un análisis de datos numéricos sobre la aparición del tumor, por ejemplo, sintomatología, tasa de mortalidad o tratamiento tras diagnóstico.
   * Información sobre sistemas de detección y diagnóstico que hayan sido desarrollados hasta la fecha y puedan ser encontrados en distintas publicaciones. Estos sistemas no deben ser necesariamente para el diagnóstico del Glioblastoma multiforme. Diferentes métricas y parámetros de evaluación del sistema inteligente que se encuentran en la literatura deberán de ser analizados para poder aplicar estos al desarrollo del sistema propio y realizar una evaluación correcta de este.
2. Obtención del conjunto de datos con los datos poblacionales y las imágenes médicas necesarias para el desarrollo del proyecto.
   * Selección de un conjunto de datos que contenga un mínimo de 50 sujetos, incluyendo imágenes y biomarcadores asociados a los sujetos.
3. Análisis del problema y planteamiento de su resolución de forma clara. Definir los diferentes pasos por los que pasarán los datos para la obtención del sistema inteligente, nombrando todas las herramientas que vayan a usarse en el proyecto y el papel que tienen estas en la resolución de este.
4. Aprendizaje del lenguaje de programación Python para su empleo en el desarrollo del proyecto.
   * Aprendizaje de Python básico.
   * Aprendizaje de las librerías de Python para el procesamiento de imagen, Machine Learning y Deep Learning.
5. Análisis de las imágenes médicas para la detección del Glioblastoma.
   * Procesamiento previo de las imágenes médicas para facilitar el paso posterior de identificación del tumor.
   * Localización del tumor mediante técnicas de procesamiento de imagen.
   * Extracción de marcadores a partir de la imagen procesada para el sistema de predicción.
   * Se plantea la detección de la zona del tumor mediante Deep Learning (identificación en la propia imagen de la zona del Glioblastoma) como mejora para el sistema. (Extra)
6. Desarrollo del sistema de predicción mediante los datos poblacionales y los marcadores extraídos mediante Machine Learning / Deep Learning.
   * Selección de los parámetros poblacionales para el desarrollo del sistema.
   * Organización y procesamiento de los datos que van a ser empleados por el sistema inteligente.
   * Elaboración del sistema de detección y predicción del desarrollo del Glioblastoma en un futuro.
7. Validación del sistema inteligente mediante el cálculo de distintas métricas de evaluación extraídas en el estado del arte. La fiabilidad del resultado final deberá ser evaluada de forma adecuada acorde a los métodos empleados en otras soluciones.

La siguiente fase se plantea como una fase opcional, la cual será desarrollada solamente en caso de contar con tiempo extra para el desarrollo del proyecto.

1. Desarrollo de una interfaz gráfica para la gestión de pacientes y exhibir los diagnósticos y las predicciones del sistema, además de datos estadísticos.
   * Definición de una BBDD para el almacenamiento de los datos de los pacientes y su gestión.
   * Interfaz gráfica que permita la gestión de los pacientes, permitiendo realizar las funciones CRUD (*Create, Read, Update and Delete*).
   * Visualización de distintas variables estadísticas.

## Planificación

El desarrollo del proyecto se va a realizar en un tiempo aproximado de 4 meses, dividido en bloques de dos meses. Aunque esto sea así, el desarrollo del proyecto también se plantea en paralelo al desarrollo de las clases. Por esta razón, se ha desarrollado un único diagrama de Gantt para la planificación, teniendo la posibilidad de poder adelantar o retrasar tareas en función del ritmo de trabajo.

Además de las fases técnicas definidas, cada uno de los bloques de dos meses tiene que ir acompañado de una memoria que recoja información sobre las tareas realizadas en ese periodo de tiempo y una presentación al finalizar el segundo bloque de tiempo dedicado al proyecto. Estos se definen como entregas o hitos, y su desarrollo deberá de darse en paralelo al desarrollo del proyecto.

De esta forma, se expone el diagrama de Gantt desarrollado en la Figura 8.



Figura 8. Diagrama de Gantt del proyecto.

# DISEÑO DE LA SOLUCIÓN

## Requisitos técnicos

Para definir los servicios que va a realizar el sistema se han definido una serie de requisitos funcionales con los que debe cumplir.

* + El sistema permitirá la visualización de las imágenes médicas y los biomarcadores de los pacientes almacenados en la BBDD.
  + El sistema permitirá la introducción de nuevos pacientes (imágenes y datos poblacionales) y se guardarán los nuevos datos en la BBDD.
  + El sistema realizará una detección del Glioblastoma a partir de las imágenes introducidas.
  + Los resultados obtenidos de la detección deben ser mostrados, incluyendo las imágenes y datos necesarios para la comprensión del resultado obtenido de forma sencilla.
  + Los resultados anteriores permitirán la confirmación del diagnóstico del paciente por parte del profesional médico correspondiente.
  + Tras la confirmación de un profesional, el modelo será reentrenado usando los datos del paciente introducido.
  + El sistema realizará una predicción del desarrollo del Glioblastoma a partir de imágenes y biomarcadores cuando se confirme el diagnóstico. Así, indicará si se prevé que el paciente sobreviva a los siguientes 12 meses o no.
  + El sistema permitirá generar un informe en un documento de texto a partir de los datos del paciente.
  + El sistema mostrará datos estadísticos que puedan ser obtenidos a partir de los datos de la BBDD.

Se han definido una serie de requisitos del software para la organización y desarrollo de distintos apartados del proyecto.

* + El sistema será desarrollado en el lenguaje de programación Python, siguiendo el estándar PEP8 para el estilo del código.
  + El código será analizado mediante SonarQube para el cálculo de distintas métricas que determinen la calidad del código y que se cumplen los estándares definidos en PEP8.
  + Git se empleará como herramienta de control de versiones, almacenando el repositorio de forma remota mediante GitHub.
  + La base de datos se definirá en SQL.

## Herramientas

# DESARROLLO

# BIBLIOGRAFÍA

[1] Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, y The National Academies of Sciences, Engineering, and Medicine, *Improving Diagnosis in Health Care*. Washington, D.C.: National Academies Press, 2015, p. 21794.

[2] H. Singh, A. Meyer, y E. Thomas, «The frequency of diagnostic errors in outpatient care: Estimations from three large observational studies involving US adult populations», *BMJ Qual. Saf.*, vol. 23, abr. 2014, doi: 10.1136/bmjqs-2013-002627.

[3] «¿Qué es el cáncer? - Instituto Nacional del Cáncer», sep. 02, 2015. https://www.cancer.gov/espanol/cancer/naturaleza/que-es (accedido nov. 13, 2020).

[4] F. Hanif, K. Muzaffar, kahkashan Perveen, S. Malhi, y S. Simjee, «Glioblastoma Multiforme: A Review of its Epidemiology and Pathogenesis through Clinical Presentation and Treatment», *Asian Pac. J. Cancer Prev.*, vol. 18, n.o 1, ene. 2017, doi: 10.22034/APJCP.2017.18.1.3.

[5] T. C. Booth, M. Williams, A. Luis, J. Cardoso, K. Ashkan, y H. Shuaib, «Machine learning and glioma imaging biomarkers», *Clin. Radiol.*, vol. 75, n.o 1, pp. 20-32, ene. 2020, doi: 10.1016/j.crad.2019.07.001.

[6] D. Armocida, A. Pesce, F. Di Giammarco, A. Frati, A. Santoro, y M. Salvati, «Long Term Survival in Patients Suffering from Glio-blastoma Multiforme: A Single-Center Observational Cohort Study», *Diagnostics*, vol. 9, n.o 4, p. 209, nov. 2019, doi: 10.3390/diagnostics9040209.

[7] C. Okuma y R. Fernández, «EVALUACIÓN DE GLIOMAS POR TÉCNICAS AVANZADAS DE RESONANCIA MAGNÉTICA», *Rev. Médica Clínica Las Condes*, vol. 28, n.o 3, pp. 360-377, may 2017, doi: 10.1016/j.rmclc.2017.05.005.

[8] «Early Detection Can Be Key to Surviving a Brain Tumor», *Weill Cornell Brain and Spine Center*, ago. 08, 2016. https://weillcornellbrainandspine.org/early-detection-can-be-key-surviving-brain-tumor (accedido dic. 09, 2020).

[9] P. J. Kelly, «Gliomas: Survival, origin and early detection», *Surg. Neurol. Int.*, vol. 1, dic. 2010, doi: 10.4103/2152-7806.74243.

[10] «Glioblastoma (GBM) – American Brain Tumor Association». https://www.abta.org/tumor\_types/glioblastoma-gbm/ (accedido dic. 09, 2020).

[11] «Early Detection», *Glioblastoma Foundation*. https://glioblastomafoundation.org/research/early-detection (accedido nov. 13, 2020).

[12] «Assessing Concordance With Watson for Oncology, a Cognitive Computing Decision Support System for Colon Cancer Treatment in Korea | JCO Clinical Cancer Informatics». https://ascopubs.org/doi/full/10.1200/CCI.17.00109 (accedido dic. 10, 2020).

[13] «Ascent of machine learning in medicine», *Nat. Mater.*, vol. 18, n.o 5, Art. n.o 5, may 2019, doi: 10.1038/s41563-019-0360-1.

[14] K. Kourou, T. P. Exarchos, K. P. Exarchos, M. V. Karamouzis, y D. I. Fotiadis, «Machine learning applications in cancer prognosis and prediction», *Comput. Struct. Biotechnol. J.*, vol. 13, pp. 8-17, ene. 2015, doi: 10.1016/j.csbj.2014.11.005.

[15] J. A. M. Sidey-Gibbons y C. J. Sidey-Gibbons, «Machine learning in medicine: a practical introduction», *BMC Med. Res. Methodol.*, vol. 19, n.o 1, p. 64, dic. 2019, doi: 10.1186/s12874-019-0681-4.

[16] S. Anto, «Supervised Machine Learning Approaches for Medical Data Set Classification-A Review», ene. 2011.

[17] S. Damilola, «A Review of Unsupervised Artificial Neural Networks with Applications», *Int. J. Comput. Appl.*, vol. 181, n.o 40, pp. 22-26, feb. 2019, doi: 10.5120/ijca2019918425.

[18] S. Mutuvi, «Introduction to Machine Learning Model Evaluation», *Medium*, feb. 05, 2020. https://heartbeat.fritz.ai/introduction-to-machine-learning-model-evaluation-fa859e1b2d7f (accedido dic. 14, 2020).

[19] K. Nighania, «Various ways to evaluate a machine learning models performance», *Medium*, ene. 30, 2019. https://towardsdatascience.com/various-ways-to-evaluate-a-machine-learning-models-performance-230449055f15 (accedido dic. 14, 2020).

[20] D. Faggella, «Machine Learning for Medical Diagnostics - 4 Current Applications», *Emerj*. https://emerj.com/ai-sector-overviews/machine-learning-medical-diagnostics-4-current-applications/ (accedido dic. 10, 2020).

[21] «Babylon Health UK - The Online Doctor and Prescription Services App», *Babylon Health*. https://www.babylonhealth.com/ (accedido dic. 11, 2020).

[22] «Better Together». https://hms.harvard.edu/news/better-together (accedido dic. 11, 2020).

[23] «Face2Gene Technology - How It Works», *Face2Gene*. https://www.face2gene.com/technology-facial-recognition-feature-detection-phenotype-analysis/ (accedido dic. 11, 2020).

[24] A. Esteva *et al.*, «Dermatologist-level classification of skin cancer with deep neural networks», *Nature*, vol. 542, n.o 7639, Art. n.o 7639, feb. 2017, doi: 10.1038/nature21056.

[25] Y. Wan, R. Rahmat, y S. J. Price, «Deep learning for glioblastoma segmentation using preoperative magnetic resonance imaging identifies volumetric features associated with survival», *Acta Neurochir. (Wien)*, vol. 162, n.o 12, pp. 3067-3080, dic. 2020, doi: 10.1007/s00701-020-04483-7.

[26] J. Valdebenito y F. Medina, «Machine learning approaches to study glioblastoma: A review of the last decade of applications», *Cancer Rep.*, vol. 2, n.o 6, p. e1226, 2019, doi: https://doi.org/10.1002/cnr2.1226.

[27] H. Sotoudeh *et al.*, «Artificial Intelligence in the Management of Glioma: Era of Personalized Medicine», *Front. Oncol.*, vol. 9, 2019, doi: 10.3389/fonc.2019.00768.

[28] Z. Qian *et al.*, «Differentiation of glioblastoma from solitary brain metastases using radiomic machine-learning classifiers», *Cancer Lett.*, vol. 451, pp. 128-135, jun. 2019, doi: 10.1016/j.canlet.2019.02.054.

[29] C. Chen, X. Ou, J. Wang, W. Guo, y X. Ma, «Radiomics-Based Machine Learning in Differentiation Between Glioblastoma and Metastatic Brain Tumors», *Front. Oncol.*, vol. 9, 2019, doi: 10.3389/fonc.2019.00806.

[30] S. Bae *et al.*, «Robust performance of deep learning for distinguishing glioblastoma from single brain metastasis using radiomic features: model development and validation», *Sci. Rep.*, vol. 10, n.o 1, Art. n.o 1, jul. 2020, doi: 10.1038/s41598-020-68980-6.

[31] L. Macyszyn *et al.*, «Imaging patterns predict patient survival and molecular subtype in glioblastoma via machine learning techniques», *Neuro-Oncol.*, vol. 18, n.o 3, pp. 417-425, jun. 2015, doi: 10.1093/neuonc/nov127.

[32] J. Jeong *et al.*, «Machine-learning based classification of glioblastoma using delta-radiomic features derived from dynamic susceptibility contrast enhanced magnetic resonance images», *Quant. Imaging Med. Surg.*, vol. 9, n.o 7, pp. 1201-1213, jul. 2019, doi: 10.21037/qims.2019.07.01.

[33] J. Lao *et al.*, «A Deep Learning-Based Radiomics Model for Prediction of Survival in Glioblastoma Multiforme», *Sci. Rep.*, vol. 7, n.o 1, Art. n.o 1, sep. 2017, doi: 10.1038/s41598-017-10649-8.