



ESCUELA DE INGENIERÍA
DE FUENLABRADA

MODELO PREDICTIVO DE LA RESPUESTA AL TRATAMIENTO
DE QUIMIO-RADIOTERAPIA EN CÁNCER DE PULMÓN
DE CÉLULAS NO PEQUEÑAS LOCALMENTE AVANZADO
BASADO EN VARIABLES RADIÓMICAS DE CT

TRABAJO FIN DE GRADO

GRADO EN INGENIERÍA BIOMÉDICA

AUTOR: Sara Altuna Pérez

TUTOR: Ángel Torrado Carvajal

CO-TUTOR: Jesús Javier Martín Pinacho

Curso académico: 2022-2023

Agradecimientos

A mi hermano. Gracias Rubén por enseñarme tu amor por la ingeniería. Contigo al lado siento que puedo llegar a donde quiera.

A mis padres, aunque nunca lo leerán. Gracias a vosotros me he convertido en la persona que siempre quise. Vuestro apoyo en cada paso que he dado durante esta etapa y todas las anteriores ha sido lo que me ha impulsado a seguir. Y así será en las que aún quedan por delante.

A mis amigas, por ser mi fuente de inspiración y hacer que quiera siempre superarme. Por motivarme cuando menos ánimos tenía. Cada trozo de este trabajo es gracias a vosotras.

Resumen

En la actualidad, el tratamiento para el cáncer de pulmón de células no pequeñas (CPCNP) localmente avanzado e irresecable está basado en la combinación de quimioterapia y radioterapia con intención radical. A pesar de que la inmunoterapia y otros tratamientos novedosos han conseguido mejorar los resultados, aumentando el porcentaje de supervivencia, menos de un tercio de los pacientes están libres de sufrir una recaída pasados cinco años. Del mismo modo, todavía no se dispone de herramientas lo suficientemente precisas y rigurosas para definir los posibles pacientes que consigan responder al tratamiento, por lo que ninguno está exento de recibir un tratamiento con un alto nivel de toxicidad, efectos adversos graves y un seguimiento uniforme independientemente de la probabilidad de respuesta.

Durante los últimos años, el ámbito de la radiómica se ha explotado notablemente y ha evolucionado concediendo la capacidad de extraer mayor cantidad de información sobre la base de las técnicas existentes de imagen a partir de datos cuantitativos. Este hecho convierte este nuevo campo en una herramienta enormemente provechosa y útil para la predicción de respuesta, así como el hallazgo de supervivientes a largo plazo permitiendo evolucionar hacia una medicina personalizada.

El propósito de este proyecto ha sido desarrollar una herramienta que combine la radiómica y las técnicas basadas en machine learning (ML) aplicándolas a la predicción de la respuesta tumoral de los pacientes incluidos en el estudio. Tras el proceso de selección de imágenes y su posterior segmentación manual, se han extraído un total de 1562 características radiómicas de distintas clases y naturaleza en base a la imagen original, que después han sido utilizadas en un completo proceso de clasificación binaria. Los resultados más relevantes obtenidos son de un 0.75 de sensibilidad, un 0.60 de especificidad, un 0.71 para el f1-score y, finalmente, un valor del 0.70 de AUC.

Para concluir el trabajo, con técnicas de explicabilidad se ha construido una manera más clara de entender el tipo de características radiómicas que más han influido en el aprendizaje de los modelos adoptados. Esta línea de investigación supone una visión prometedora en la predicción de supervivencia de los pacientes oncológicos, aún con un gran recorrido de crecimiento por delante.

Palabras Clave: Cáncer de pulmón, inteligencia artificial, radiómica, python, TAC.

Índice general

| | |
|---|------------|
| Índice de figuras | v |
| Índice de cuadros | vii |
| Lista de Acrónimos y Abreviaciones | x |
| 1. Introducción | 1 |
| 1.1. Anatomía del Pulmón | 1 |
| 1.2. El Cáncer de Pulmón y sus Tipos | 2 |
| 1.3. Epidemiología del Cáncer de Pulmón | 3 |
| 1.4. Pronóstico del Cáncer de Pulmón | 5 |
| 1.5. Causas y Factores de Riesgo del Cáncer de Pulmón | 5 |
| 1.6. Diagnóstico del Cáncer de Pulmón | 6 |
| 1.7. Tratamiento del Cáncer de Pulmón | 7 |
| 1.7.1. Cirugía | 8 |
| 1.7.2. Radioterapia | 9 |
| 1.7.3. Quimioterapia | 10 |
| 1.7.4. Inmunoterapia | 11 |
| 1.7.5. Terapia dirigida | 11 |
| 1.8. Principios de la Tomografía Axial Computarizada | 11 |
| 2. Marco Teórico y Estado del Arte | 15 |

| | |
|---|-----------|
| 2.1. El Aprendizaje Automático | 15 |
| 2.2. Radiómica | 17 |
| 2.3. Estado del Arte | 18 |
| 3. Motivación y Objetivos | 21 |
| 3.1. Motivación | 21 |
| 3.2. Hipótesis | 22 |
| 3.3. Objetivos | 22 |
| 4. Materiales | 25 |
| 4.1. Descripción de la Base de Datos | 25 |
| 4.1.1. Sujetos de Estudio | 25 |
| 4.1.2. Tipo de Imágenes | 26 |
| 4.2. Herramientas Utilizadas | 28 |
| 4.2.1. 3D Slicer | 28 |
| 4.2.2. Python | 28 |
| 5. Metodología | 29 |
| 5.1. Selección de Imágenes | 29 |
| 5.2. Segmentación de los Tumores | 29 |
| 5.2.1. Creación de las Máscaras | 31 |
| 5.3. Extracción de Características Radiómicas | 31 |
| 5.3.1. Selección de Características | 34 |
| 5.3.2. Estandarización de las Características | 37 |

| | |
|--|-----------|
| 5.4. Arquitectura de Clasificación | 37 |
| 5.4.1. Optimización de los Hiperparámetros | 40 |
| 5.5. Evaluación del Modelo | 44 |
| 5.6. Explicabilidad del Modelo | 47 |
| 6. Resultados | 49 |
| 6.1. Rendimiento del Clasificador | 49 |
| 6.2. Estudio de las Características | 50 |
| 7. Discusión | 53 |
| 7.1. Impacto de la Segmentación | 54 |
| 7.2. Rendimiento del Clasificador | 55 |
| 7.3. Estudio de las Características | 57 |
| 8. Conclusión y Trabajos Futuros | 61 |
| 8.1. Conclusión | 61 |
| 8.2. Trabajos Futuros | 62 |
| Bibliografía | 65 |

Índice de figuras

| | |
|---|----|
| 1.1. Anatomía de los pulmones [2]. | 1 |
| 1.2. Evolución anual de la mortalidad por CP en hombres de España y la Comunidad de Madrid [8]. | 4 |
| 1.3. Evolución anual de la mortalidad por CP en mujeres de España y la Comunidad de Madrid [8]. | 4 |
| 1.4. Tipos de cirugía de pulmón [15]. | 9 |
| 1.5. Esquema de funcionamiento de un TAC [24]. | 13 |
| 3.1. Diagrama de Gantt sobre la planificación temporal del proyecto. | 23 |
| 4.1. Ejemplo de una imagen de TAC. | 27 |
| 4.2. Ejemplo de una imagen de TAC con el tumor marcado en rojo. | 27 |
| 5.1. Ejemplos de los tumores segmentados con 3D Slicer | 30 |
| 5.2. Esquema de la metodología aplicada desde la recolección de datos hasta la extracción de características radiómicas. | 34 |
| 5.3. Ejemplo de una matriz de correlación obtenida sobre las características originales de primer grado. | 35 |
| 5.4. Ejemplo de una matriz de correlación tras eliminar las características originales de primer orden cuyo índice de Pearson es mayor que 0.8. | 36 |
| 5.5. Esquema del funcionamiento interno de un algoritmo <i>Random Forest</i> [51]. | 38 |
| 5.6. Diagrama de validación cruzada para $K = 10$ [54]. | 40 |
| 5.7. Ejemplo de los mejores valores de los hiperparámetros. | 43 |
| 5.8. Curva ROC-AUC en una situación ideal de clasificación [60]. | 46 |

| | |
|---|----|
| 5.9. Curva ROC-AUC en una situación real de clasificación [60]. | 46 |
| 5.10. Esquema del flujo de trabajo de SHAP [61]. | 47 |
| 6.1. Curva ROC obtenida tras el entrenamiento del Random Forest. | 50 |
| 6.2. Gráfico con la descripción global de las características calculada como la suma de todos los valores absolutos de SHAP para cada una de las variables. | 51 |
| 6.3. Histogramas de las cuatro características más importantes según el algoritmo SHAP. | 52 |

Índice de cuadros

| | |
|---|----|
| 1.1. Incidencia del cáncer de pulmón en el mundo y en España [6]. | 3 |
| 5.1. Rango de valores por hiperparámetro. | 41 |
| 5.2. 10 mejores realizaciones llevadas a cabo en la búsqueda aleatoria. | 41 |
| 5.3. Valores elegidos de cada hiperparámetro. | 42 |
| 5.4. Valores óptimos introducidos en el Random Forest de cada hiperparámetro. . . | 42 |
| 6.1. Resultados de la clasificación empleada. | 49 |

Lista de Acrónimos y Abreviaciones

- CADe** Sistemas de Apoyo a la Detección
- CADx** Sistemas de Apoyo al Diagnóstico
- CNN** Red Neuronal Convolutiva
- CNPT** Condiciones Normales de Presión y Temperatura
- CPCNP** Cáncer de Pulmón de Células No Pequeñas
- CPCP** Cáncer de Pulmón de Células Pequeñas
- CP** Cáncer de Pulmón
- DICOM** Digital Imaging and Communication In Medicine
- DL** Deep Learning
- FDG** Fluorodesoxiglucosa
- FN** Falso Negativo
- FP** Falso Positivo
- HURYC** Hospital Universitario Ramón y Cajal
- HU** Unidades Hounsfield
- IARC** International Agency for Research on Cancer
- IA** Inteligencia Artificial
- ML** Machine Learning
- NIFTI** Neuroimaging Informatics Technology Initiative
- NRRD** Nearly Raw Raster Data Format
- PACS** Sistema de Comunicación y Archivado de Imágenes
- PET** Tomografía por Emisión de Positrones
- REDECAN** Red Española de Registros de Cáncer
- RM** Resonancia Magnética
- ROI** Region of Interest/Región de Interés

LISTA DE ACRÓNIMOS Y ABREVIACIONES

- SAI** Sistema de Alimentación Ininterrumpida
- SBRT** Radioterapia estereotáctica
- SEOM** Sociedad Española de Oncología Médica
- SHAP** Shapley Additive Explanations
- SNC** Sistema Nervioso Central
- STL** Standard Triangle Language
- SVM** Support Vector Machine
- TAC** Tomografía Axial Computarizada
- VN** Verdadero Negativo
- VP** Verdadero Positivo

Capítulo 1

Introducción

1.1 Anatomía del Pulmón

Los pulmones son los órganos que se encargan de llevar a cabo la función respiratoria. Se ubican en el tórax a ambos lados del corazón y están protegidos por las costillas. En la parte más craneal del aparato respiratorio se encuentra la tráquea, a partir de la cual nacen los bronquios. Ambos se dividen en dos ramas que entran en los pulmones. Finalmente, de estos últimos salen diversas bifurcaciones denominadas bronquiolos. Como se ve en la Figura 1.1, el pulmón derecho tiene un tamaño mayor ya que cuenta con tres lóbulos mientras que el izquierdo tiene únicamente dos. A su vez, cada lóbulo se divide en pequeños segmentos a los que llegan los bronquiolos, que se subdividen terminando en alvéolos donde, finalmente, se realiza el intercambio gaseoso [1].

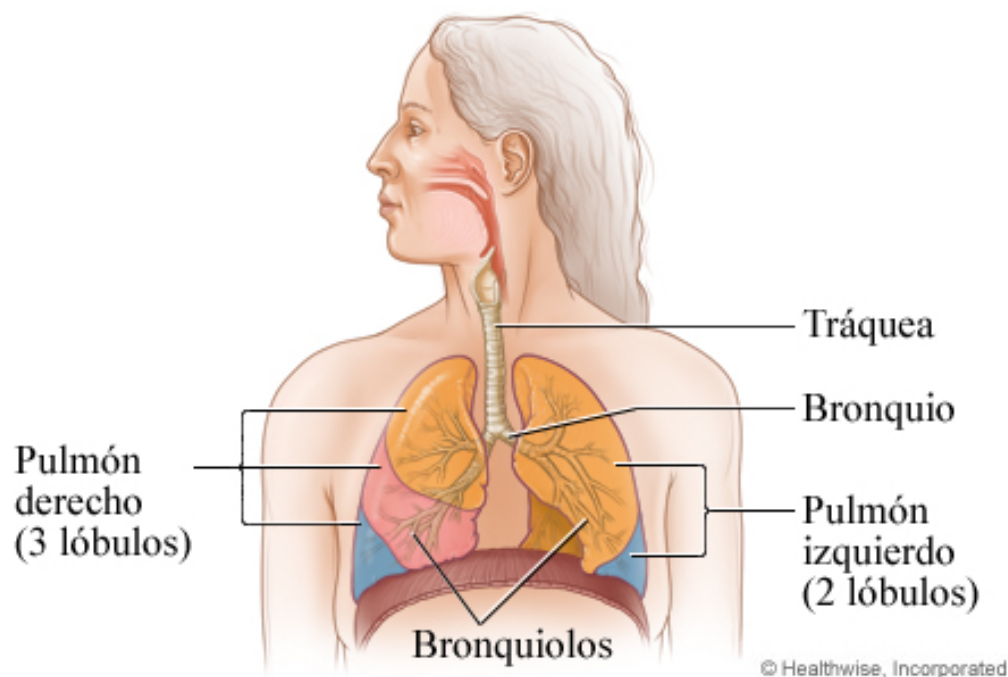


Figura 1.1: Anatomía de los pulmones [2].

Recubriendo los pulmones se encuentra la pleura, una membrana serosa que evita el roce con la pared interna de la cavidad torácica. Esta se compone de dos capas, la parietal en la parte externa y la visceral en la parte interna. Entre ambas capas hay un espacio denominado cavidad pleural que, normalmente, está lleno de un líquido lubricante. Además, debajo de los pulmones se halla un músculo plano en forma de cúpula llamado diafragma, que disminuye el espacio en la cavidad torácica durante la inhalación y lo incrementa durante la exhalación [1].

1.2 El Cáncer de Pulmón y sus Tipos

El cáncer de pulmón (CP) es una patología formada en los tejidos pulmonares, en general en aquellas células que cubren los conductos de aire: bronquios, bronquiolos o alveolos pulmonares. Concretamente, se trata de una multiplicación y división descontrolada de dichas células formando una masa denominada tumor, nódulo o lesión. El problema radica en la alta capacidad que posee para ocupar otros tejidos y órganos. Estas células patológicas pueden transportarse por medio de la sangre o a través del líquido linfático que se encuentra rodeando el tejido pulmonar. La linfa se mueve por los vasos linfáticos vaciándose en los ganglios linfáticos, unos órganos pequeños que protegen contra las infecciones. El flujo de salida habitual de este líquido en los pulmones es hacia dentro del tórax. Cuando una célula cancerosa es movida hacia uno de estos ganglios linfáticos o hacia otra parte del cuerpo por medio de los vasos sanguíneos, el cáncer comienza a extenderse poco a poco, dando lugar a la metástasis [3].

Existen dos tipos principales de CP, el cáncer de pulmón de células pequeñas (CPCP), cuya característica principal es el tamaño reducido de las células a las que afecta; y el cáncer de pulmón de células no pequeñas (CPCNP), que se trata de células de un tamaño mayor a las anteriores. Cada uno de ellos tiene un desarrollo y tratamiento diferentes. El CPCNP es el más común y en el que se centra este proyecto. Para determinar el tratamiento, es primordial conocer la clase de CPCNP que tiene un paciente. Existen tres tipos principales [4]:

- Carcinomas de células escamosas: Se localiza en las células escamocelulares, que recubren el interior pulmonar. El 30% de los CPCNP son de células escamosas.
- Adenocarcinomas: En este caso, se inicia en las células epiteliales que revisten el pulmón por la parte exterior. Es el tipo más frecuente abarcando un 40% de todos los casos de CPCNP.

- Carcinomas de células grandes: Están ubicadas en cualquier parte del pulmón. Es por eso que no se clasifican como carcinoma ni como adenocarcinoma de células escamosas.

1.3 Epidemiología del Cáncer de Pulmón

De acuerdo con los datos de la International Agency for Research on Cancer (IARC), el CP es el tipo de patología oncológica con mayor mortalidad (1.589.925 muertes) e incidencia (1.824.701 casos nuevos) en todo el mundo y el segundo cáncer más diagnosticado después del cáncer de próstata. Concretamente, el CPCNP constituye entre el 80% y el 85% de todos los diagnósticos de cáncer de pulmón, siendo más frecuente en hombres (1.241.601 casos) que en mujeres (583.100 casos) [5].

En España, se estima que en 2022 se han diagnosticado 30.948 casos nuevos de CP, entre ellos, aproximadamente 22.300 son hombres y el resto mujeres, una incidencia elevada en comparación con el resto de países europeos. Por grupos de edad, las personas con más de 75 años suponen un 32% de los nuevos casos, al igual que el grupo de 65 a 74 años. Por otro lado, las personas con edades comprendidas entre los 55 y los 64 años constituyen el 25% seguidos de los menores de 54 hasta los 15 años con un 11% de la incidencia [6].

| Incidencia Anual | Número de Casos | Hombres | Mujeres |
|------------------|-----------------|-----------|---------|
| En el Mundo | 1.824.701 | 1.241.601 | 583.100 |
| En España | 30.984 | 22.300 | 8.648 |

Cuadro 1.1: Incidencia del cáncer de pulmón en el mundo y en España [6].

La mortalidad del CP ha ido disminuyendo poco a poco en los últimos años debido a los avances tecnológicos como el desarrollo de nuevas herramientas y tratamientos y mejora de equipos de diagnóstico. Un convenio entre la Sociedad Española de Oncología Médica (SOEM) y la Red Española de Registros de Cáncer (REDECAN), revela que la supervivencia del CP en 5 años ha aumentado como consecuencia de este progreso desde los años 2002-2007 hasta el periodo del 2008-2013, pasando del 11,2% al 12,7% en hombres [7]. En la Figura 1.2 se puede ver esta evolución. Sin embargo, como consecuencia del aumento del consumo de tabaco en mujeres a partir de la década de los 70, el desarrollo de la mortalidad ha sido muy desfavorable. En la Figura 1.3 vemos este cambio en la población femenina desde 1975 hasta 2017.

1.3. EPIDEMIOLOGÍA DEL CÁNCER DE PULMÓN

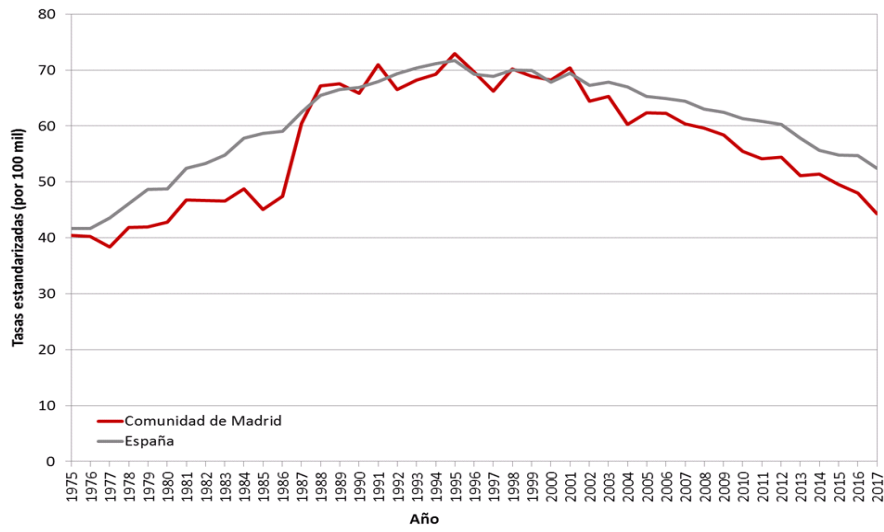


Figura 1.2: Evolución anual de la mortalidad por CP en hombres de España y la Comunidad de Madrid [8].

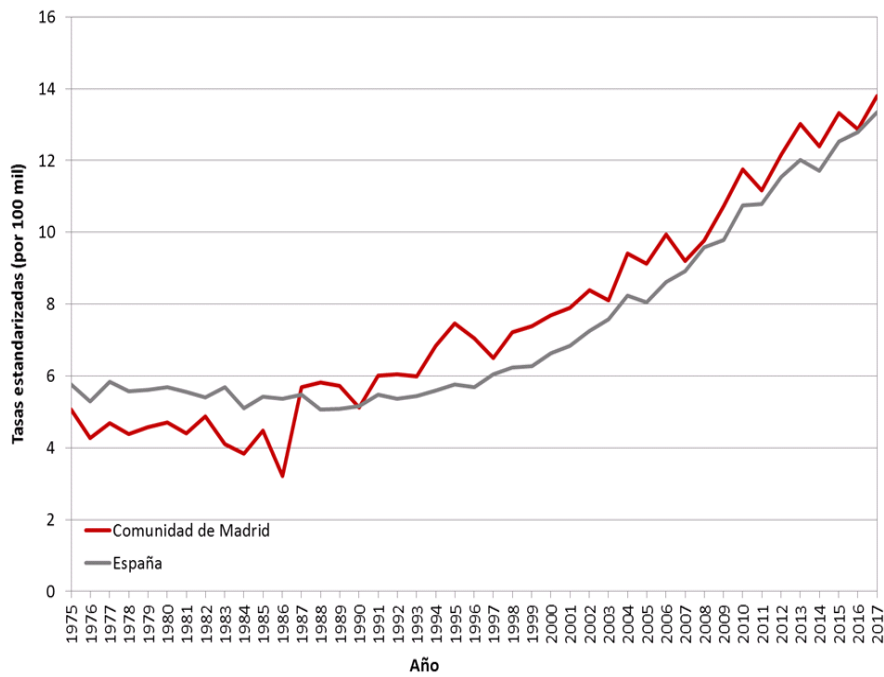


Figura 1.3: Evolución anual de la mortalidad por CP en mujeres de España y la Comunidad de Madrid [8].

1.4 Pronóstico del Cáncer de Pulmón

El pronóstico del CP va a depender del estadio en el que se encuentra la enfermedad en el momento del diagnóstico. Con frecuencia, los pacientes son valorados y diagnosticados cuando la enfermedad está en una fase avanzada ya que habitualmente es complicado observar los síntomas en fases tempranas. De la misma manera, en muchas ocasiones la enfermedad es diagnosticada en pacientes de altas edades con otros tipos de enfermedades cuyos síntomas hacen dificultoso el tratamiento oncológico.

Otro factor importante a tener en cuenta para evaluar el pronóstico es la herencia genética. Específicamente, en el cáncer pulmonar, los tumores hereditarios son inusuales, pero existe un pequeño porcentaje de los pacientes diagnosticados que se encuentran en esta situación. Entre el 5% y el 10% de los casos de CP son relacionados con mutaciones de los genes adquiridas durante la vida, como fumar o vivir en lugares con una alta contaminación. Es primordial tener en cuenta que el CP no se hereda, sino que se desarrolla un mayor riesgo a padecerlo. Algunas investigaciones recientes y ensayos clínicos buscan mejorar la supervivencia de los sujetos que padecen este tipo de cáncer de manera que en los próximos años se produzca un incremento notable tanto en la calidad de vida de los pacientes con CP como en la supervivencia de este [9].

1.5 Causas y Factores de Riesgo del Cáncer de Pulmón

La principal causa del CP es el tabaquismo. Es por ello que una bajada del consumo de tabaco implica un descenso de la incidencia del CP, aunque es necesario un tiempo de unos años de latencia para disminuir el riesgo de padecer la enfermedad al de una persona no fumadora. Según la estadística, alrededor de un 80% y un 90% de los casos de CP ocurren en fumadores o en ex fumadores, aunque también se diagnostica en no fumadores pero con unas características especiales [10].

Existen otros factores que incrementan el riesgo de desarrollar cáncer pulmonar, como la edad avanzada o el sexo masculino. Es probable que esta última característica esté altamente relacio-

nada con el mayor hábito tabáquico en hombres que en mujeres. Sin embargo, este hecho está cambiando en los últimos años observando un aumento de la incidencia entre el sexo femenino. Asimismo, existen otras causas que pueden contribuir al avance de un tumor pulmonar [11]:

- Polución en zonas urbanas.
- Exposición a sustancias como el radón o el asbesto en el área de trabajo.
- Antecedentes familiares de CP.

1.6 Diagnóstico del Cáncer de Pulmón

El diagnóstico del CP proporciona una forma saber el tipo de cáncer (si es CPCP o CPCNP), pero también permite conocer el estadio en el que se encuentra (del 0 al IV). Con el fin de conseguir el tratamiento óptimo, es necesario llevar a cabo un diagnóstico adecuado.

En primer lugar, es imprescindible una anamnesis y exploración física apropiadas para conocer factores importantes de la vida cotidiana como el tabaquismo o el trabajo y otros datos como la posible presencia de adenopatías, fatiga, tos persistente o dolores. De manera complementaria, existen otras técnicas como el diagnóstico por imagen. Hay cuatro principales y las más usadas para este tipo de patología [12]:

- Radiografía de tórax: Es la primera prueba que se realiza cuando existe una sospecha de CP. Sirve para asegurar el diagnóstico y resulta de vital importancia para una detección precoz del CP en los pacientes con factores de riesgo. La ventaja es que tiene una duración y un coste muy bajos.
- Tomografía Axial Computarizada (TAC): Consiste en un estudio radiológico con el que se obtienen imágenes del cuerpo en tres dimensiones. Es muy útil para diferenciar entre los tipos histológicos de CP aunque no permite discriminar con fiabilidad entre ellos. La duración de esta prueba es de 10-15 minutos y es necesario un tamaño del tumor mayor de 8 mm para poder llevar a cabo esta técnica.
- Tomografía por Emisión de Positrones (PET): Se trata de una prueba de imagen tridimensional que permite detectar las zonas del organismo del paciente con actividad gracias a

sustancias marcadas radioactivamente distribuidas por el cuerpo como la glucosa. Normalmente, los cánceres consumen glucosa a una velocidad mayor que el resto de tejidos no patológicos por lo que es posible detectar aquellas zonas del organismo en las que existen células tumorales. El mayor inconveniente es que esto dificulta la discriminación entre tumor e inflamación o infección debido a que en ambos escenarios existe un aumento de la actividad metabólica. Además, como en la TAC, el tamaño de las lesiones también deben ser por encima de 8 mm.

- **Resonancia Magnética (RM):** Es la prueba elegida para un estudio del Sistema Nervioso Central (SNC), ya que permite conocer si el cáncer se ha extendido o no hacia otros órganos. Esta técnica se utiliza escasas veces en el caso del CP, únicamente si existe clínica neurológica y no se identifica afectación metastásica con la TAC.
- **Gammagrafía ósea:** Es una herramienta para determinar si el tumor se ha propagado a los huesos por medio de la inyección de un isótopo radiactivo que se fija al hueso dañado. En general, esta prueba no es necesaria cuando antes se ha realizado un TAC, PET o ambas combinadas.

1.7 Tratamiento del Cáncer de Pulmón

El CP puede ser tratado de diferentes maneras. La elección del tipo de tratamiento depende de distintos factores:

- Tipo de CP, es decir, si es de células pequeñas o de células no pequeñas, escamoso o no escamoso, etc.
- Estadio en el que está la enfermedad.
- Localización y tamaño del tumor.
- Peculiaridades de cada individuo (edad, comorbilidades asociadas, estado nutricional, etc.).

Para tratar el CP existen cuatro diferentes métodos terapéuticos: cirugía, radioterapia, quimioterapia, inmunoterapia y terapia dirigida. La quimioterapia junto con la terapia dirigida y la

inmunoterapia, son terapias sistémicas, es decir, los medicamentos se reparten por todo el organismo por medio de la sangre. Sin embargo, la radioterapia y la cirugía son tratamientos locales ya que se centran solo en la zona del tumor. Frecuentemente, es necesaria la combinación de varias para llevar a cabo el tratamiento [13].

1.7.1 Cirugía

Es el tratamiento habitual para personas con un estado de salud bueno y con CP localizado, esto es, cuando el tumor afecta exclusivamente al pulmón sin invadir a otros órganos u otros grandes vasos. La cirugía puede ser utilizada tanto para el tratamiento como con fines diagnósticos.

Dependiendo de la cantidad de pulmón que se extraiga, podemos diferenciar tres procedimientos quirúrgicos distintos [14]:

- **Resección sublobar:** Cuando el tamaño del tumor es muy reducido o las funciones pulmonares no admitan la extracción de un lóbulo completo, se extirpa solo una porción pequeña del tejido donde se ubica el tumor. Dentro de esta existen dos tipos que se llevarán a cabo dependiendo el paciente y donde se encuentre el tumor:
 - **Resección en cuña:** Se realiza en los tumores localizados periféricamente en pacientes excesivamente enfermas para una lobectomía.
 - **Segmentectomía:** Se elimina el segmento completo en el que se encuentra el tumor.
- **Lobectomía:** Cuando el tumor se encuentra en uno de los lóbulos del pulmón derecho o del izquierdo, se extrae la totalidad de uno de ellos y los ganglios linfáticos que haya asociados a este.
- **Pneumonectomía:** Consiste en extraer el pulmón completo afectado por el tumor. Este tipo de cirugía no está indicada para pacientes de edades avanzadas o con una disminución de las funciones pulmonares.

En la Figura 1.4 se puede observar las porciones del pulmón extirpadas en cada uno de los casos anteriormente explicados.

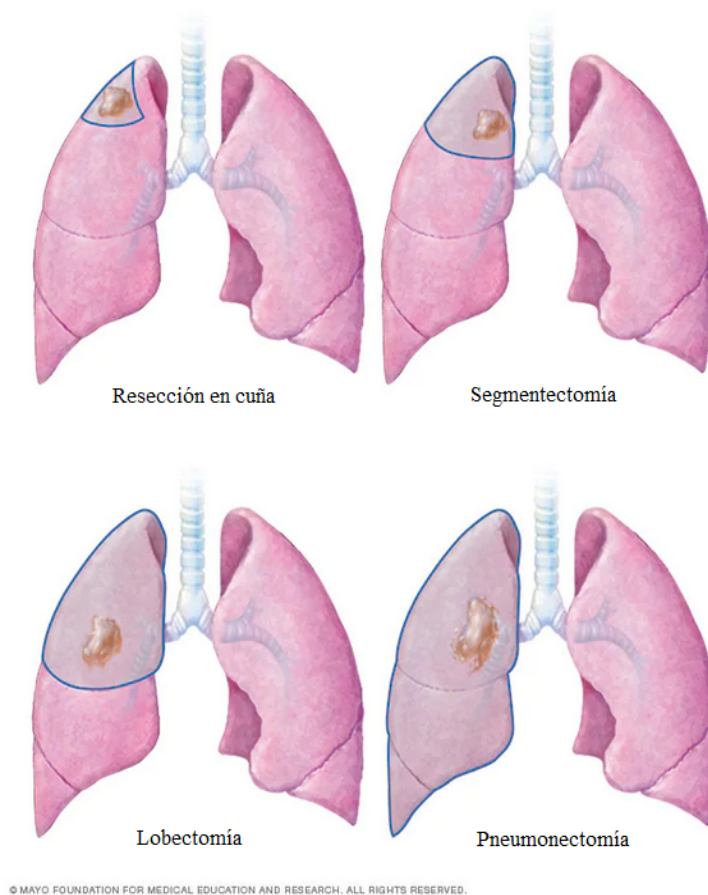


Figura 1.4: Tipos de cirugía de pulmón [15].

1.7.2 Radioterapia

La radioterapia o terapia de radiación emplea altas dosis de radiación para reducir el tumor y destruir las células cancerosas o frenar su crecimiento dañando su ADN. Cuando esto ocurre, las células paran de dividirse o mueren, descomponiéndose y finalmente desechándolas del cuerpo. La elección de este tratamiento puede ser con la intención de eliminar el tumor de forma drástica y que no vuelva a crecer, o paliativa, buscando reducir su tamaño para mejorar la calidad de vida del paciente [16].

Dependiendo de donde proceda la radiación, existen dos tipos de radioterapia: de haz externo y braquiterapia.

1. Radioterapia de haz externo: La radiación procede de un equipo externo al cuerpo que enfoca el haz en el tumor. Este tipo de radioterapia reduce el daño del tejido que rodea el nódulo debido a la disminución de la radiación.
2. Braquiterapia: La fuente de radiación tiene origen en el interior del cuerpo del paciente y esta puede ser líquida (terapia sistémica) o sólida. En este tipo de tratamiento se colocan cápsulas con radiación dentro del paciente en las zonas del tumor o cerca de este.

Si bien esta terapia consiste en aplicar una dosis de radiación determinada en el tumor y los tejidos que lo envuelven, es posible desarrollar efectos adversos más allá de la zona tumoral. En el caso del CP, la región del cuerpo radiada es el tórax, por lo que los pacientes pueden desarrollar problemas respiratorios, dolor de garganta o dificultad para tragar. Con la radioterapia esterotáctica (SRBT), cada vez estos efectos secundarios son menos lesivos ya que el depósito de radiación es más preciso en la lesión a tratar y menos en los tejidos circundantes [17].

1.7.3 Quimioterapia

La quimioterapia se basa en la administración de drogas citotóxicas muy potentes que detienen el crecimiento anómalo de las células tumorales, bien eliminándolas o bien evitando su división. En el CP, la quimioterapia es utilizada en todas las fases excepto en el estadio I del CPCNP. El suministro de estos fármacos puede llevarse a cabo por vía oral, es decir, en forma de pastillas, cápsulas, comprimidos o líquidos, o por vía intravenosa. Estos últimos son los más comunes ya que esta forma de administración permite que las sustancias vayan directamente al torrente sanguíneo. En ocasiones, algunos medicamentos intravenosos pueden encontrarse también en su versión para administración oral [18].

Debido a que la quimioterapia es un tratamiento sistémico que afecta a todas las células del cuerpo, los efectos secundarios más habituales se hallan en otros órganos y zonas del cuerpo: caída del pelo, náuseas y vómitos, pérdida de apetito o estreñimiento [19]. Para minimizarlos, la quimioterapia se aplica con pausas temporales llamadas ciclos. Generalmente, las sesiones se aplican cada tres semanas para dar tiempo a que las células sanas se regeneren.

1.7.4 Inmunoterapia

El sistema inmunitario protege al organismo de microorganismos malignos y agentes externos invasores combatiéndolos. Cuando los linfocitos detectan una célula maligna, acuden a destruirla o alertan a otro tipo de células para eliminarlas. A diferencia de las otras terapias, la inmunoterapia está dirigida a estimular este sistema inmunitario para que las células tumorales sean reconocidas como malignas. Además, se reduce la toxicidad de los fármacos ya que solo las células cancerígenas son atacadas, dejando intactas a aquellas que están sanas [20].

1.7.5 Terapia dirigida

Los tumores pueden estructurarse dependiendo de su mutación. Estas mutaciones se pueden utilizar como dianas biológicas para que los profesionales oncólogos puedan pronosticar que tipo de tratamiento va a resultar más efectivo en cada paciente. La terapia dirigida consiste en medicamentos más específicos y menos tóxicos que la quimioterapia que impiden los procedimientos que favorecen el crecimiento de las células. La mayoría de estos fármacos inhiben la formación de vasos sanguíneos o angiogénesis con el fin de parar la vascularización de los nódulos. Indirectamente, esto afecta al desarrollo de estos [21].

1.8 Principios de la Tomografía Axial Computarizada

La tomografía axial computarizada (TAC) se basa en tomar imágenes radiográficas desde distintas proyecciones del paciente para posteriormente construir una imagen tridimensional. Todos los equipos de TAC se constituyen de cinco partes principales [22]:

- El gantry: Es el cuerpo vertical del equipo que gira y que contiene un orificio central en el que se introduce la camilla del paciente. En su interior se encuentra el tubo de rayos X, el generador de alto voltaje, los detectores y los colimadores.
- La mesa del paciente: Es una cama donde se tumba al paciente durante la exploración. Se

fabrica de un material de baja absorción para que no actúe como filtro.

- El ordenador de reconstrucción: Es el cuello de botella de todos los TACs donde se obtiene la imagen tridimensional final del paciente.
- Consola de control: Es un ordenador con dos pantallas, una con la historia clínica y otros datos del paciente, y otra con la imagen en reconstrucción.
- Sistema de Alimentación Ininterrumpida (SAI): Es una fuente de alimentación que sirve para no perder las imágenes ya tomadas en el caso de que la luz del se vaya.

Desde el interior del gantry, el tubo de rayos X dispara un haz de radiación bien colimado que se dirige hacia paciente, que permanece recostado en la camilla. Cuando lo traspasa, llega al detector un coeficiente de la cantidad de atenuación que se ha producido. Finalmente, obtenemos la suma de todas las atenuaciones de los distintos tejidos por los que los fotones han pasado. Como el gantry gira alrededor del paciente, a un mismo detector le llega la información de la atenuación de distintas proyecciones. Cuantos más fotones llegan, menos atenuación habrá y por tanto más negra se ve la imagen. Por el contrario, cuando la imagen se ve blanca, significa que se han atenuado muchos fotones del haz de radiación y han llegado menos al detector. Cada vez que la fuente de rayos X hace un giro completo, el ordenador lleva a cabo una serie de operaciones matemáticas para finalmente obtener la reconstrucción bidimensional de uno de los cortes del paciente. Tras obtener toda la información de un corte completo, la imagen es almacenada y la mesa del paciente se mueve progresivamente hacia el interior del gantry. Este proceso se repite hasta conseguir el número de cortes deseado [23]. En la Figura 1.5 podemos ver un esquema del desarrollo de la técnica.

Para poder comparar los distintos tejidos mediante una escala cuantitativa, se utilizan las Unidades Hounsfield (HU), que lleva a cabo una reasignación en función de la atenuación del agua destilada en Condiciones Normales de Presión y Temperatura (CNPT).

$$HU = 1000 \times \frac{\mu_t - \mu_{agua}}{\mu_{agua} - \mu_{aire}} \quad (1.1)$$

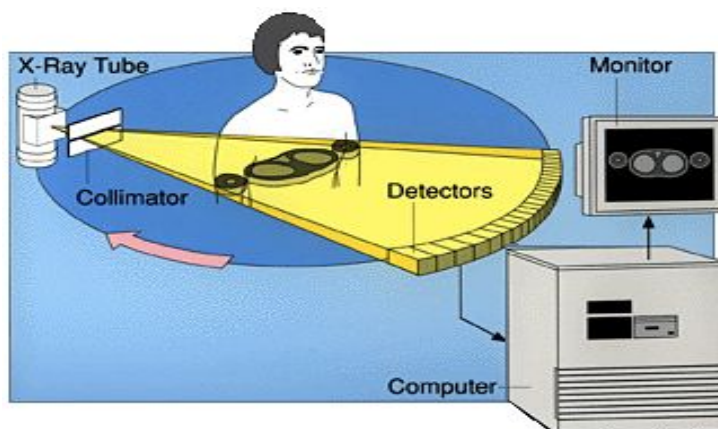


Figura 1.5: Esquema de funcionamiento de un TAC [24].

Capítulo 2

Marco Teórico y Estado del Arte

2.1 El Aprendizaje Automático

La inteligencia artificial (IA) ha ido experimentando un crecimiento exponencial a lo largo de los últimos años. La IA y otras diferentes tecnologías conforman una potente herramienta en muchos entornos cotidianos haciendo situaciones de la vida diaria más sencillas. El aprendizaje automático o machine learning (ML) es una disciplina derivada del ámbito de la IA que tiene como propósito crear sistemas capaces de imitar el comportamiento humano aprendiendo a reconocer patrones entre millones de datos. Aprender automáticamente denota la no intervención humana, es decir, que estos sistemas mejoran de manera independiente en el tiempo [25]. En el entorno médico, la IA ha supuesto una nueva manera de visualizar la medicina ya que es utilizada como apoyo a la toma de decisiones y para el análisis de imágenes médicas. Las herramientas de apoyo asisten a los profesionales sanitarios a la hora de decidir sobre diagnósticos, tratamientos, medicamentos u otras necesidades dando un rápido acceso a las investigaciones o información destacable para atender a los pacientes. En el ámbito de las imágenes médicas, la IA resulta un instrumento útil para el análisis de TAC, RM o incluso placas de rayos X y evitar pasar por alto hallazgos que el ojo humano podría omitir.

Los fundamentos básicos de las diversas técnicas de IA son aspectos matemáticos como elementos de álgebra lineal (vectores y matrices), geometría, probabilidad y estadística básica o análisis de funciones variables reales como derivadas e integrales. En resumen, la IA es construida mediante algoritmos y datos para entrenarlos. Debido a la alta variedad de tareas que el ML puede llevar a cabo, se han creado tres ramas principales: el aprendizaje supervisado, no supervisado y aprendizaje profundo.

El aprendizaje supervisado resuelve problemas conocidos y necesita datos de entrada y de salida que estén etiquetados durante el entrenamiento [26]. Existen dos tipos principales de aprendi-

zaje supervisado: la clasificación y la regresión. La clasificación consiste en la asignación de las muestras eligiendo entre dos o más clases. Para ello, se utiliza la información aprendida de los datos de entrenamiento y se elige la etiqueta correcta. Los algoritmos mas comunes son:

- Clasificador bayesiano: Se basa en métodos probabilísticos que asumen que la existencia o falta de una propiedad específica no está relacionada con cualquier otra característica.
- Support Vector Machine (SVM): Es un conjunto de algoritmos cuyo objetivo es encontrar un hiperplano que divida dos clases diferentes de la mejor manera posible.
- Árbol de decisión: Es un modelo predictivo que emplea una búsqueda exhaustiva para asociar todas las observaciones que tienen valores parecidos de la variable de respuesta.
- K vecinos más cercanos o k-NN: Utiliza una medida de cercanía para predecir la agrupación de un punto de datos individual.

A diferencia de la clasificación, la regresión se utiliza para establecer una relación entre variables dependientes e independientes a partir de un rango continuo de valores. Por tanto, la salida no es discreta, sino que se trata de una función dependiente de los parámetros de entrada. En este caso, los algoritmos más utilizados son:

- Regresión logística: Es una técnica de análisis de datos que predice una variable categórica en función de la variable independiente.
- Regresión por mínimos cuadrados: Es un método que estima los coeficientes de la regresión lineal para ajustar los datos en una recta que minimiza la suma de los cuadrados de las distancias verticales entre dichos puntos y la recta.

Por otro lado, en el aprendizaje no supervisado, a diferencia del anterior, el modelo aprende a partir de datos en los que no se conoce la etiqueta. Debido a esta ausencia de datos de salida, no puede aplicarse directamente a problemas de clasificación o regresión. Sin embargo, el objetivo principal de estos modelos es hallar la estructura que siguen los datos de entrada agrupándolos conforme a características semejantes compartidas [27].

En último lugar, se encuentra el aprendizaje profundo o deep learning (DL). El DL es completamente diferente del ML. Se trata de una disciplina que tiene en cuenta dos o más variables

creando modelos computacionales formados por varias capas de procesamiento, de forma que la red pueda construir diversos niveles de abstracción representativos del conjunto de datos [28].

2.2 Radiómica

La radiómica es una metodología en auge que estudia de forma no invasiva las características que existen en las imágenes médicas y que son invisibles al ojo humano, por medio de algoritmos automáticos [29]. La finalidad de estos procesos es asociar cada característica a estados fisiológicos específicos. Sustancialmente, la radiómica extrae datos de las imágenes por medio de métodos matemáticos. Estas características se extraen a lo largo de una región de interés con nódulos que podrían haber pasado desapercibidos.

La radiómica puede aplicarse en una gran variedad de modalidades de imágenes médicas como RM, TAC o PET. Esto permite una combinación de los resultados para obtener un planteamiento cruzado con un valor complementario. Sin embargo, si bien es cierto que el estudio de la radiómica aun está en las primeras fases de desarrollo, se prevé que en el futuro va a contribuir en gran medida a mejorar los tratamientos y profundizar en el conocimiento de la heterogeneidad de las distintas patologías que subyacen a las enfermedades.

Una de las aplicaciones más útiles de la radiómica es predecir el subtipo tumoral más probable basándose en la imagen. Frente a las muestras anatomopatológicas mediante biopsias, esta herramienta presenta la ventaja de llevar a cabo un análisis de todo el volumen tumoral sin ceñirse al cilindro de material obtenido. De esta manera, se consigue mucha más información. También es de mucha utilidad en aquellos pacientes en los que la biopsia no es una opción factible por causas como la proximidad entre el tumor y las vías respiratorias o los vasos sanguíneos grandes [30].

La radiómica no solo se centra en analizar las características del tumor sino que también estudia el área peritumoral que lo rodea. Así, se puede obtener información valiosa sobre la infiltración por células del sistema inmunológico o sobre la densidad de los vasos pequeños, dos características muy importantes para determinar la evolución del tumor [31]. Algunos estudios recientes se han centrado en la investigación sobre la posible predicción del grado de respuesta patológica del CP tras un tratamiento de quimio-radioterapia aplicando radiómica [32]. Actualmente, en la práctica clínica, dicha evaluación está basada únicamente en el tamaño del tumor eludiendo otra

información significativa como variaciones en la histología subyacente que puede ser estudiado con radiómica.

2.3 Estado del Arte

El campo de la imagen médica ha sido explotado creando algoritmos y probando nuevos métodos diferentes de clasificación de imágenes para conseguir reemplazar las técnicas invasivas que evalúan el estadio de los nódulos pulmonares. Generalmente, se lleva a cabo una comparación de diferentes métodos para ver cuál tiene un mejor comportamiento. La mayoría de las tareas que implican estadiaje y evaluación de la respuesta tumoral del CP son mediante TAC para conseguir una segmentación de los tumores y caracterizar su histología y estadio. Sin embargo, existen procedimientos con un componente invasivo.

Algunos estudios se centran en evaluar la eficacia en el diagnóstico de métodos funcionales, como la PET con Fluorodesoxiglucosa (FDG), frente a métodos de imágenes estructurales, como el TAC. En primer lugar, el PET utilizando FDG permite seguir la actividad de la glucosa para diferenciar las lesiones malignas de las benignas y, aunque resulta complementario, este método funcional es característicamente mejor que el método estructural para la estadificación torácica no invasiva del CPCNP [33]. Por otro lado, son muchos los estudios publicados acerca de la posible redundancia al utilizar métodos de PET/TC integrados. La precisión de este procedimiento es demasiado baja para sustituir la estadificación invasiva de los ganglios linfáticos torácicos en pacientes con CPCNP. Aunque siempre se consigue un mayor beneficio y precisión cuando se hace un uso de las dos en conjunto que por separado [34].

Las neoplasias pulmonares malignas se han caracterizado considerablemente mediante el empleo de técnicas basadas en radiómica y DL. Algunos se han utilizado para segmentación automática de los órganos en riesgo para la planificación de la radioterapia en CP, o con el objetivo de identificar a los mejores candidatos para someterse a un tratamiento de inmunoterapia o de terapia dirigida. En el futuro, se espera que la radiómica influya en mayor medida en el tratamiento de los tumores pulmonares para mejorar y optimizar el proceso desde el principio del diagnóstico hasta el final del seguimiento y pasando por el tratamiento completo [35].

Como consecuencia de la capacidad de las características radiómicas de describir la huella genética del tumor altamente relacionada con la agresividad de éste y su posible curación, es posible

utilizarlas con el propósito de diseñar modelos para predecir el resultado del tratamiento del cáncer realizado (quimioterapia, radioterapia, inmunoterapia, terapia dirigida) o una combinación de ellas [36]. Además, en este contexto, se han desarrollado otros modelos en el ámbito de la radiómica para vaticinar la posible metástasis a distancia en pacientes con CPCNP usando características del TC o de este combinado con imágenes PET [37].

La histología de los tumores constituye un predictor relevante para los resultados terapéuticos del CP. El método más confiable de clasificación histológica es tomar una pequeña muestra del tejido para revisión patológica. No obstante, el DL se ha desarrollado en el análisis de las imágenes médicas de forma que se puedan utilizar los datos radiológicos para describir detalladamente las características de los tejidos cancerígenos. El estudio llevado a cabo por Tafadzwa L Chaunzwa [38] hace uso de una Red Neuronal Convolutiva (CNN) para predecir la histología de lesiones en CP. Se utilizó una red neuronal VGG-16 en la que se entrenó un amplio conjunto de imágenes de TAC de pacientes con adenocarcinoma y carcinoma de células escamosas, optimizando los hiperparámetros de forma iterativa. Los resultados muestran una alta efectividad de las CNN en tareas de distinción entre los tipos histológicos más habituales de CPCNP, siendo capaz de detectar adenocarcinoma con una especificidad mayor que la sensibilidad. Esto representa un punto de partida prometedor de los estudios de CP con DL.

Capítulo 3

Motivación y Objetivos

3.1 Motivación

Este proyecto consiste en la extracción de características radiómicas y entrenamiento de algoritmos de aprendizaje automático para desarrollar herramientas de apoyo a la decisión médica en el CP motivado por varias razones. Se ha llevado a cabo en colaboración con el Hospital Universitario Ramón y Cajal (HURYC) [39], desde donde se ha propuesto la necesidad a cubrir, se han obtenido las imágenes del conjunto de datos y se ha supervisado la evolución del TFG.

El aumento ocurrido durante los últimos años de la prevalencia de los pacientes con CP hace que el CP ocupe la primera causa de muertes oncológicas en el mundo y el segundo lugar en incidencia mundial. Por ello, es necesario estimular la búsqueda y desarrollar nuevas técnicas para cambiar la forma de enfocar los tratamientos disponibles y conseguir un progreso en la obtención de resultados en estos pacientes. Esto convierte al diagnóstico precoz de la enfermedad en uno de los puntos más importantes en la práctica clínica diaria. Las imágenes radiológicas, y en concreto sus características extraíbles, son un instrumento potente no invasivo para la detección del CP y su diagnóstico y futuro tratamiento. Éstas, en combinación con técnicas de radiómica, suponen un nuevo enfoque de sistemas de apoyo a la detección (CADe) y sistemas de apoyo al diagnóstico (CADx), sirviendo de ayuda a los médicos en lugar de sustituirlos. De esta manera, se contribuye a mejorar en gran medida la supervivencia de estos pacientes ya que los sistemas informáticos de los que se disponen en la actualidad poseen la capacidad de percibir información que el ojo humano sería incapaz. Por consiguiente, con el análisis de estas imágenes médicas se pueden extraer nuevos datos ocultos con un valor clínico muy elevado, evaluando fácil y rápidamente la presencia o el crecimiento de un tumor durante el tratamiento.

Otra de las motivaciones más importantes es la extrapolación de los resultados a otras patologías como el cáncer de mama, pudiendo usar los algoritmos para el mismo fin que este estudio

o incluso para detección de ganglios centinela. En último lugar, con este proyecto se pretende lograr una implementación real en la clínica para ayudar a los médicos especialistas y conseguir mejorar la calidad de vida de los pacientes oncológicos.

3.2 Hipótesis

La hipótesis del proyecto es: *Las características radiómicas obtenidas del análisis tumoral basado en la imagen de TAC pueden aplicarse para predecir el grado de respuesta del tumor a un tratamiento de quimio-radioterapia y su posibilidad de progresión en el futuro.*

3.3 Objetivos

El objetivo primario del proyecto es desarrollar y evaluar un modelo predictivo de respuesta al tratamiento de los CPCNP en estadio III a partir de las características radiómicas del tumor extraídas de la imagen de TAC.

Para ello, es necesaria una lista de objetivos secundarios:

- Recoger y anonimizar imágenes de TAC de tórax de pacientes diagnosticados con CP.
- Segmentar manualmente los tumores a partir de las imágenes seleccionadas.
- Extraer características radiómicas de los nódulos segmentados y analizar su correlación con el subtipo histológico.
- Estudiar los cambios en las características radiómicas de la masa tumoral post-tratamiento para predecir la persistencia de la enfermedad.
- Seleccionar las características más relevantes para dotar de sentido al algoritmo de ML.
- Valorar y discutir los resultados obtenidos.

En la Figura 3.1 se recoge una estimación del tiempo invertido en cada uno de los objetivos mencionados:

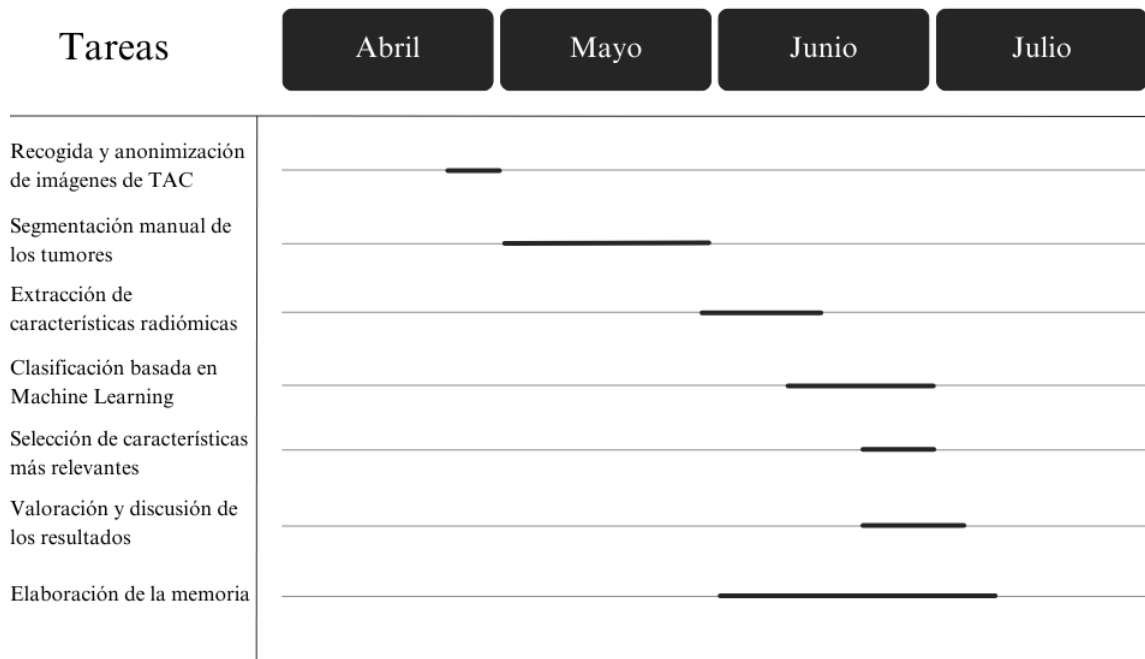


Figura 3.1: Diagrama de Gantt sobre la planificación temporal del proyecto.

Capítulo 4

Materiales

En esta sección presentaremos las herramientas utilizadas para el desarrollo del proyecto y cómo se ha llegado hasta ellas. En concreto, se explica el proceso llevado a cabo desde la elección de los sujetos que conformarán la base de datos, así como sus imágenes, hasta los softwares que se han manejado para conseguirlas y llevar a cabo la extracción de sus características radiómicas.

4.1 Descripción de la Base de Datos

4.1.1 Sujetos de Estudio

Las imágenes utilizadas en el proyecto han sido proporcionadas por el Servicio de Radiología Cardiorádica del HURYC [40].

El proyecto se ha desarrollado siguiendo la normativa de Protección de Datos de Carácter Personal. En primer lugar, todos los datos recogidos se identifican únicamente con códigos, de forma que solo las personas autorizadas puedan relacionarlos con la historia clínica del paciente. Así, la identidad de estos no podrá ser revelada excepto si se tratase de los profesionales sanitarios e investigadores, y solo bajo situaciones de urgencia médica. Por otro lado, los Comités de Ética de la Investigación únicamente podrán acceder para llevar a cabo una investigación con el fin de comprobar el procedimiento del estudio clínico y los datos personales de los pacientes. También si el propósito es asegurar el cumplimiento de las normas de buena práctica clínica. Además, no se revelará la identidad de los sujetos si se publican los resultados del estudio. Así pues, se conservará en todo momento la privacidad de la información extraída.

La realización del estudio se lleva a cabo sin la obtención del consentimiento informado de los pacientes ya que esto supondría un esfuerzo desproporcionado pudiendo obstaculizar en gran

medida el éxito del proyecto. Las razones principales son:

- El estudio es de carácter retrospectivo y realizado con imágenes ya existentes almacenadas en el Sistema de Archivado y Comunicación de Imágenes (PACS) del Servicio de Radiodiagnóstico.
- Los datos utilizados están completamente anonimizados y no contienen identificadores que pudieran dar lugar a relaciones con las personas de las que provienen.
- El proyecto realizado no constituye un riesgo para la vida de los pacientes que participan en el mismo.
- Se respetará la solicitud expresa de un paciente de no ser incluido en este proyecto.
- Según lo indicado en el Artículo 14, apartado 5b, del Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016, referente a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos, el proyecto puede llevarse a cabo sin la obtención del consentimiento informado puesto que los datos se utilizarán con fines de investigación científica [41].

4.1.2 Tipo de Imágenes

El conjunto de datos consiste en imágenes de pacientes diagnosticados de CPCNP localmente avanzado irresecable y tratados con quimioterapia y radioterapia con intención radical.

La población consta inicialmente de un total de 67 pacientes, 50 hombres y 17 mujeres, con diagnóstico de CP confirmado, todos mayores de edad con edades comprendidas entre los 45 y 85 años. Las imágenes utilizadas para la extracción de características radiómicas son imágenes de TAC de tórax con un grosor de corte de más de 3 mm y con contraste intravenoso tomadas en el plano axial. Se han descartado pacientes cuyos TACs no cumplan con estos requisitos o con imágenes realizadas más de 8 semanas después de haber terminado el tratamiento. También aquellos que no lo hayan completado en su totalidad por cualquier motivo.

En la Figura 4.1 vemos un ejemplo de imagen utilizada en una zona sin tumor. Por otro lado, la Figura 4.2 muestra un corte de la imagen donde se observa como se ve un fragmento del tórax con nódulo.



Figura 4.1: Ejemplo de una imagen de TAC.



Figura 4.2: Ejemplo de una imagen de TAC con el tumor marcado en rojo.

4.2 Herramientas Utilizadas

4.2.1 3D Slicer

Tras descargar todas las imágenes anonimizadas del software Synapse 3D utilizado en el Servicio de Radiología del HURYC, se cargaron en formato DICOM en 3D Slicer, un software gratuito de código abierto utilizado para visualizar, procesar, segmentar, registrar y analizar imágenes en tres dimensiones [42]. En el caso de este proyecto, se ha utilizado el programa para segmentar todos los tumores de cada una de las imágenes en el plano axial, utilizando el módulo Segmentations. Se van cargando las imágenes una a una, buscando la zona tumoral y segmentándola manualmente. Finalmente, se ha guardado en una carpeta individual para cada paciente la segmentación correspondiente con la imagen DICOM.

4.2.2 Python

Python es un lenguaje de programación extensamente manejado en ML en el que se pueden llevar a cabo tareas de orientación a objetos, programación imperativa y funcional. Es fácil de aprender, eficiente y ejecutable en una gran variedad de plataformas distintas [43]. En este proyecto se ha utilizado la versión 3.6.13. Su uso es gratuito ya que los derechos de autor pertenecen a una fundación sin ánimo de lucro.

Py-Radiomics

Dado que el objetivo de este proyecto es extraer características radiómicas que caractericen el tumor, se ha utilizado el paquete *Py-Radiomics*, una herramienta sencilla de Python de código abierto que se utiliza para extraer características de radiómica de imágenes médicas que esclarezcan de forma cuantitativa las segmentaciones llevadas a cabo [44]. Para hacer uso de esta herramienta se ha utilizado Jupyter [45], que permite la identificación fácil y rápida de posibles errores y la visualización de los resultados obtenidos.

Capítulo 5

Metodología

En este capítulo vamos a describir detalladamente y de forma ordenada los pasos que se han seguido para llevar a cabo la investigación. En primer lugar, cómo se han seleccionado las imágenes tumorales y la posterior segmentación y creación de sus máscaras correspondientes. También se explicará el proceso de extracción de las características radiómicas y la mejor selección de un subconjunto de estas. Finalmente, especificaremos el modelo elegido para llevar a cabo la tarea de clasificación, su evaluación con las métricas más útiles y la explicabilidad del mismo.

5.1 Selección de Imágenes

Se hizo un estudio observacional retrospectivo unicéntrico de los pacientes que se añadieron a la base de datos del HURYC, que contiene las imágenes de TAC de todos los pacientes diagnosticados de CP. En primer lugar se escogieron los pacientes que padecían CPCNP y se excluyeron los que tenían CPCP. Después, se revisó cuáles de estos no habían completado el tratamiento en su totalidad y se descartaron. Finalmente, se analizaron los TACs de tórax de cada uno de ellos y se eligieron a los que tenían el grosor de corte y contraste requeridos. Cuando se consiguió un estudio en el que todos los sujetos cumplieran los criterios de inclusión, se procedió a anonimizar sus imágenes para comenzar con las descargas.

5.2 Segmentación de los Tumores

Se ha llevado a cabo una segmentación manual imagen por imagen de todos los tumores gracias al módulo *Segmentations* de 3D Slicer. El primer paso es cargar la imagen DICOM del

primer paciente y hacer una búsqueda de todos los cortes del TAC en los que existe tumor. Una vez localizada la zona tumoral se comienza a segmentar. Para ello, se delimita el contorno del nódulo en cada uno de los cortes con la herramienta *Draw*. Como resultado, se obtiene una representación en tres dimensiones del tumor. En la Figura 5.1 vemos cuatro ejemplos de ello.

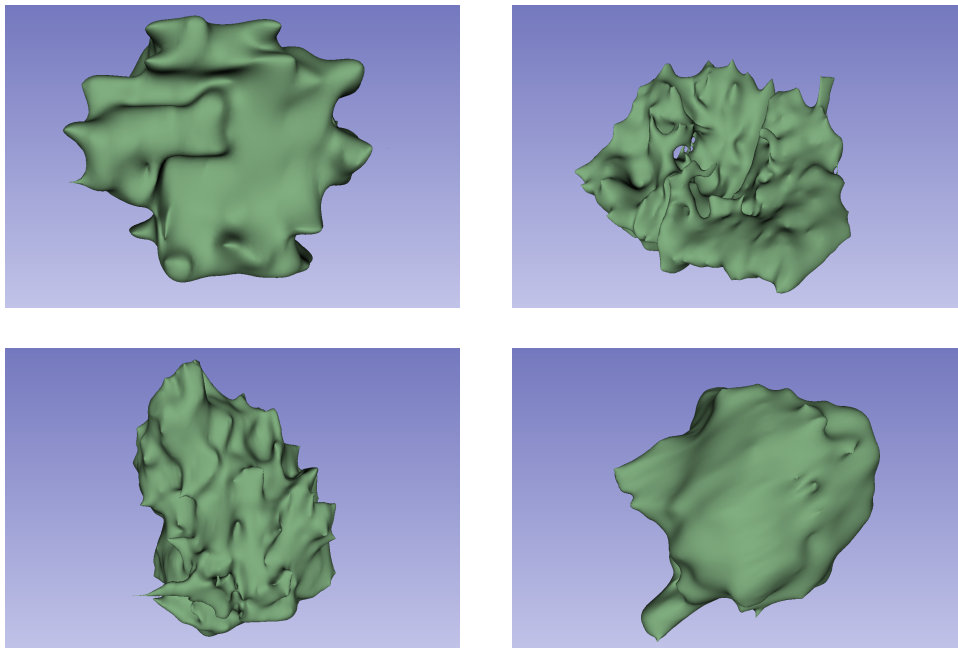


Figura 5.1: Ejemplos de los tumores segmentados con 3D Slicer

El proceso se repite con todas las imágenes de todos los pacientes de la base de datos obteniendo un total de 67 tumores segmentados. Todas las segmentaciones se guardaron inicialmente en formato Nearly Raw Raster Data (NRRD) para obtener una versión editable. El siguiente paso fue la revisión por el Doctor Jesús Javier Martín Pinacho, radiólogo cardiotorácico del HURYC y co-tutor de este TFG, de todas las segmentaciones realizadas y, tras obtener su validación, se convirtieron al formato de estereolitografía (STL) ya definitivo para la posterior creación de sus correspondientes máscaras binarias.

5.2.1 Creación de las Máscaras

En este paso se elabora la máscara de cada una de las segmentaciones, que no es más que una representación también en tres dimensiones del tumor en relación con la imagen del TAC correspondiente. Ambos archivos, tanto la máscara como la imagen radiológica, se guardarán en el formato Neuroimaging Informatics Technology Initiative (NIfTI).

Con toda esta información, se elabora una carpeta por cada uno de los sujetos que llamamos `sub_XXX` siendo `XXX` el número de paciente comprendido entre 001 y 067. En cada una de éstas se almacena la máscara de la segmentación del nódulo (`XXX_mask.nii.gz`) junto con la imagen original de TAC asociada a dicho tumor (`XXX_img.nii.gz`).

5.3 Extracción de Características Radiómicas

El módulo de extracción de características utiliza lenguaje Python. Py-Radiomics puede usarse de distintos modos. En el caso de este proyecto, se ha utilizado el módulo en Modo de Parámetros. Esto significa que se hará un uso de un fichero llamado *Param.yaml* con unos determinados filtros y características relevantes seleccionados previamente. En este modo será necesario dicho archivo para extraer características, pero también la imagen, la segmentación y la máscara. El fichero está dividido en tres partes:

- *setting*: Fija los parámetros que pueden ser utilizados para el pre-procesamiento de las imágenes y la extracción de sus características. Debido a que las imágenes que se utilizan en este proyecto son extraídas directamente de la base de datos del HURyC, es innecesario indicar los parámetros relacionados con el pre-procesamiento de estas.
- *imageType*: Define los filtros que se quieren aplicar a la imagen.
- *featureClass*: Determina las características que deben ser extraídas.

Py-Radiomics tiene abundantes características que se basan en el histograma de las imágenes, la intensidad, características morfológicas o sobre la textura. A continuación, se hace una breve descripción de las ocho clases en las que se engloban [46]:

- **First Order Features:** Describen la distribución de las intensidades de los vóxeles dentro de la región de la imagen definida por la máscara. Algunas de estas características son: energía, entropía o los percentiles.
- **Shape Features (3D):** Incluyen descriptores de forma y tamaño tridimensional de la ROI. Algunos ejemplos son el volumen de vóxeles o la esfericidad.
- **Shape Features (2D):** En este caso, se engloban descriptores de forma y tamaño bidimensional de la ROI, Entre ellas se encuentra la superficie de los vóxeles o el perímetro.
- **Gray Level Co-occurrence Matrix (GLCM) Features:** Describe una función de probabilidad conjunta de segundo orden de una región restringida por la máscara. En esta clase se comprenden características como la autocorrelación o el promedio conjunto.
- **Gray Level Size Zone Matrix (GLSZM) Features:** Cuantifica las zonas de nivel de gris de una imagen. Por ejemplo: la falta de uniformidad del nivel de gris o la variación de nivel de gris.
- **Gray Level Run Length Matrix (GLRLM) Features:** Esta matriz cuantifica las ejecuciones de nivel de gris, es decir, la cantidad de píxeles consecutivos que tienen el mismo valor de gris en unidades de longitud. Se incluyen características como el énfasis a corto y a largo plazo o el porcentaje de ejecución.
- **Neighbouring Gray Tone Difference Matrix (NGTDM) Features:** Cuantifica la diferencia entre el valor de gris de un píxel con el promedio de los valores de gris de los píxeles de alrededor. Algunos ejemplos son el contraste, la fuerza o la complejidad.
- **Gray Level Dependence Matrix (GLDM) Features:** Cuantifica la dependencia que existe entre niveles de gris en la imagen. Entre estas características se incluye el énfasis en dependencia pequeña o el gran énfasis en la dependencia.

Todas estas características se pueden calcular sobre la imagen original o a partir de una imagen derivada a la cual se le fijan los filtros definidos en el fichero, exceptuando las características de forma (Shape 3D o Shape 2D) ya que estas no son calculadas a partir de la intensidad de los píxeles de la imagen, sino que se extraen a partir de la ROI de la máscara. En este proyecto se han considerado todas las clases de características en todos los tipos de imágenes disponibles en Py-Radiomics, entre ellas:

- Original: La imagen sin filtros.
- Wavelet: Se llevan a cabo todas las combinaciones posibles de filtro paso alto y paso bajo en las tres dimensiones disponibles, de forma que se obtienen 8 descomposiciones por cada nivel de la imagen.
- LoG: Es el filtro laplaciano de Gauss que sirve para realzar los bordes.
- Square: Calcula el cuadrado del conjunto de intensidades de la imagen y las escala linealmente con respecto del rango original.
- SquareRoot: Calcula la raíz cuadrada del conjunto de intensidades de la imagen y las reduce con respecto del rango original.
- Logarithm: Calcula el logaritmo de la intensidad absoluta +1.
- Exponential: Calcula la exponencial filtrando el valor de intensidad absoluta.
- Gradient: Devuelve el valor del gradiente local.
- LocalBinaryPattern2D: Toma el patrón binario local en una operación bidimensional.
- LocalBinaryPattern3D: Toma el patrón binario local en tres dimensiones mediante armónicos esféricos.

El siguiente paso es la implantación del extractor de características. En la Figura 5.2 se puede ver un esquema del procedimiento seguido desde la adquisición de las imágenes de los pacientes hasta la extracción todas las características de cada uno de los pacientes. Adicionalmente, se creó una nueva columna correspondiente a la etiqueta de cada sujeto: 1 si responde al tratamiento y, por el contrario, 0 si no lo hace. Finalmente se ha obtenido una base de datos con 1563 columnas, una por cada característica (originales y derivadas de la imagen) más la columna de etiquetas, y 67 filas, una por cada paciente.

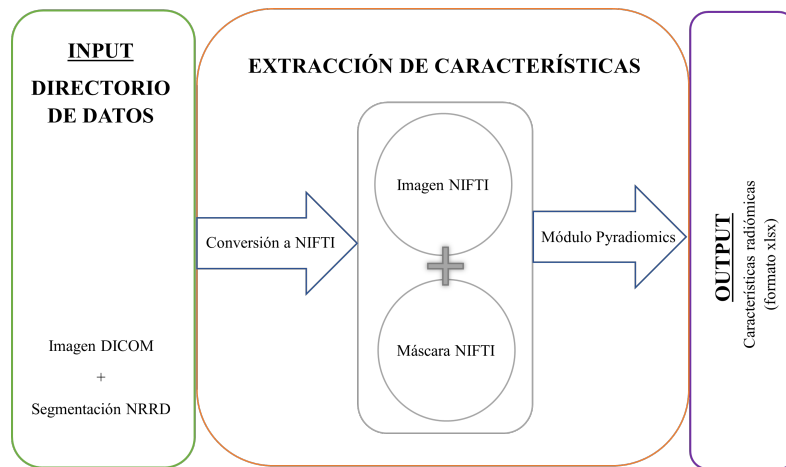


Figura 5.2: Esquema de la metodología aplicada desde la recolección de datos hasta la extracción de características radiómicas.

5.3.1 Selección de Características

Es posible que muchas de las características extraídas sean de poco valor, redundantes o incluso constantes, que solo aporten ruido al modelo. El primer tipo de características que podemos desestimar son las constantes. Esta clase de variables no aportan información para resolver el problema que abordamos puesto que el valor que contienen es el mismo para los pacientes que responden al tratamiento y para los que no. Tras examinar el conjunto de datos obtenido, se han detectado varias variables constantes que han sido eliminadas de la base de datos.

Una vez hecho esto, se ha pasado a analizar la correlación entre las características dentro de cada grupo de estas. La correlación es una medida que indica cómo de relacionadas están dos variables y, de ser así, cuál es el valor de la fuerza de esa relación. Para conseguir una aproximación visual de este concepto se ha utilizado la matriz de correlación. Se trata de una matriz que enfrenta todas las variables con todas en una tabla de doble entrada donde a cada una de ellas se le asigna un valor que fluctúa de -1 a 1, llamado coeficiente de correlación de Pearson [47], ρ_{xy} . Este coeficiente mide la dependencia o independencia lineal entre dos variables cuantitativas donde 1 o -1 significa una dependencia completa directa e inversa, respectivamente, y 0 representa la ausencia de una relación lineal de las dos variables analizadas [48]. De esta forma, cuando se enfrenta una característica consigo misma, obtenemos un valor de 1 o -1. Este coeficiente viene dado por la siguiente expresión:

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y} \tag{5.1}$$

donde σ_{xy} es la covarianza entre las dos variables y el denominador representa el producto entre sus desviaciones típicas.

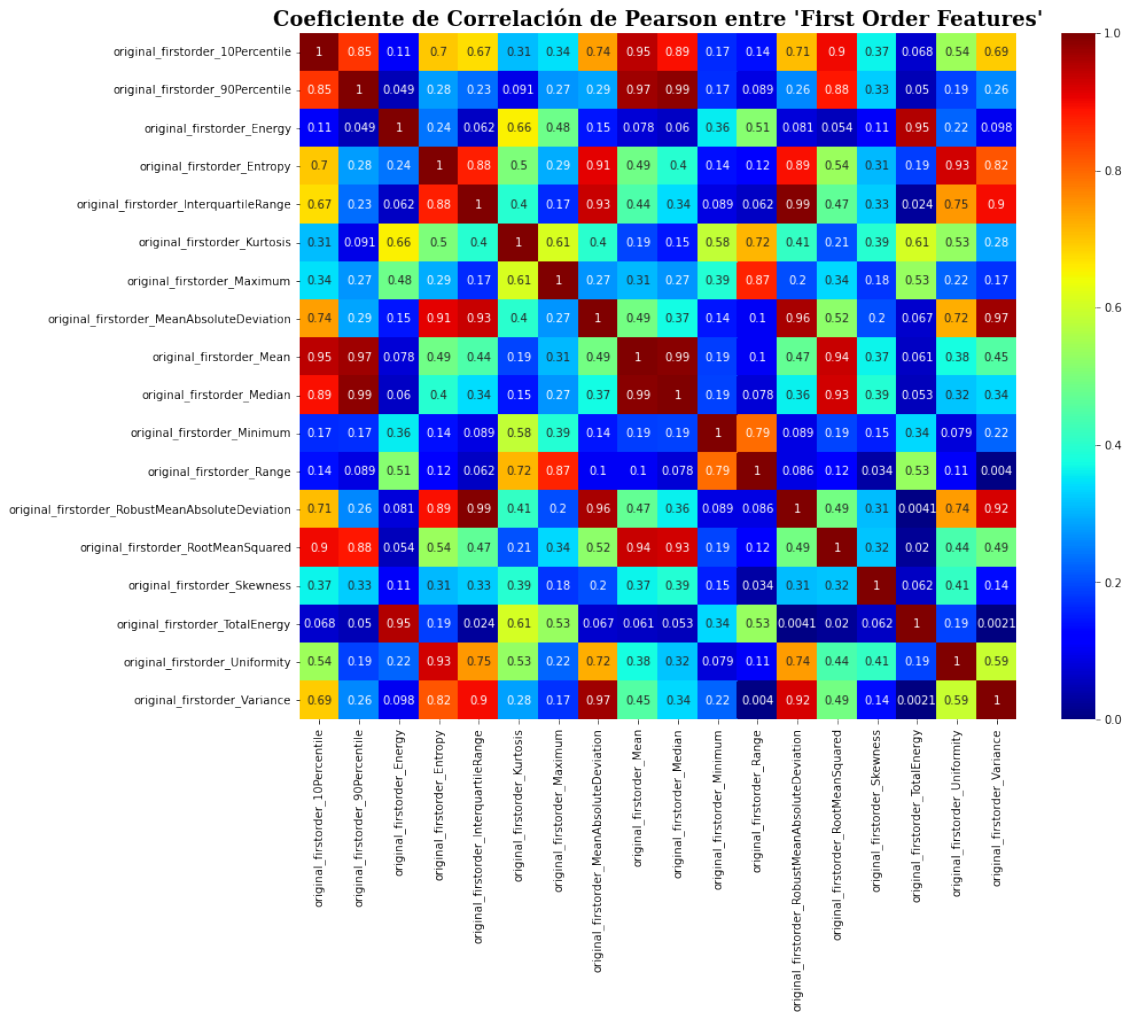


Figura 5.3: Ejemplo de una matriz de correlación obtenida sobre las características originales de primer grado.

Observando la Figura 5.3, podemos ver gran cantidad de valores del índice de correlación por encima de 0.8. Esto indica que aquellas características con este perfil se encuentran muy correlacionadas entre ellas, por lo que la información que aportan es prácticamente la misma y podemos prescindir de una de ellas. Un ejemplo claro es la característica *Energy*, cuyo índice

5.3. EXTRACCIÓN DE CARACTERÍSTICAS RADIÓMICAS

de correlación con respecto de la característica *Total Energy* es de 0.95. Sin duda alguna, podemos descartar una de las dos. Tras analizar todas las variables del grupo de la misma manera, obtenemos una nueva matriz cuyas características ya no están relacionadas unas con otras, como se muestra en la Figura 5.4.

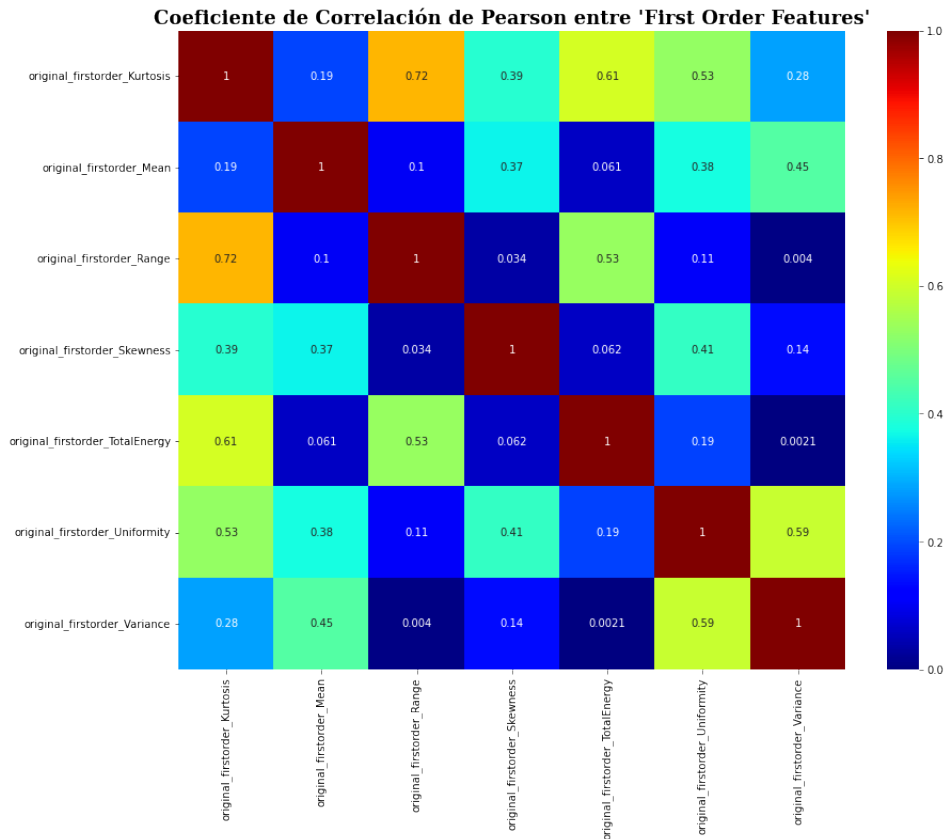


Figura 5.4: Ejemplo de una matriz de correlación tras eliminar las características originales de primer orden cuyo índice de Pearson es mayor que 0.8.

Este proceso se repite con todos los grupos de características originales (Shape, GLCM, etc.) y derivadas (Wavelet, Square, Exponential, etc.), reduciendo la dimensionalidad a 532 características.

Sin embargo, este número sigue siendo demasiado elevado en comparación con la pequeña cantidad de pacientes que existen. Para mejorar aún más el entrenamiento y la precisión del modelo de clasificación futuro, implantaremos un algoritmo de selección de características llamado *SelectKBest* [49]. Este método proporciona una selección de las K mejores variables independientes que consiguen una puntuación más alta. Para nuestro problema hemos utilizado

la función *f_classif*, una métrica basada en la técnica de análisis de varianza (ANOVA). Este cálculo trata de encontrar un valor para cada característica que indica la capacidad de discriminación entre las dos clases de cada una de ellas. En este caso, hemos fijado el valor de K a 30, por lo que, como resultado, el algoritmo nos devolverá el nombre de las 30 características que tengan un mayor valor de la puntuación elegida. Estas son las características finales con las que se entrenará el clasificador.

5.3.2 Estandarización de las Características

Gran cantidad de algoritmos de aprendizaje automático tienen un rendimiento mejor cuando las variables se encuentran en una escala parecida y cercana a una distribución normal. Concretamente, la estandarización consiste en variar los valores de modo que la desviación estándar de la distribución sea 1. Implícitamente se lleva a cabo un escalado, que equivale a convertir el rango de los valores de las características. Existen varios métodos para ejecutarlo, en nuestro caso se ha utilizado la función *StandardScaler* de *Scikit-Learn*. Este algoritmo resta la media de todos los valores de una característica y después se divide entre la desviación estándar. El resultado es una distribución con media aproximada de 0 y una desviación estándar de 1.

5.4 Arquitectura de Clasificación

Para el proceso de clasificación del proyecto, se ha introducido un método de aprendizaje supervisado basado en árboles de decisión, llamado *Random Forest* [50]. Consiste en la unión de diferentes árboles de decisión en vez de utilizar solamente uno, formando un bosque que actúan como un conjunto. Cada uno de esos árboles tiene un aspecto de árbol invertido con nodos raíz, nodos internos y nodos hoja y clasificará de una manera diferente. Al terminar el proceso, la decisión que más veces se haya obtenido será la que finalmente va a predecir el algoritmo (Ver Figura 5.5).

La idea principal de cualquier *Random Forest* (RF) es sencilla pero muy vigorosa: el conocimiento de la multitud. Un alto número de árboles individuales no correlacionados entre ellos es la clave de este modelo. De esta forma el clasificador será capaz de predecir en conjunto de manera más precisa que cada árbol individualmente. La explicación de este suceso es que cada

árbol ampara a los demás de los errores individuales que cada uno de ellos pueda producir. Si bien unos árboles harán predicciones incorrectas, otros habrán acertado.

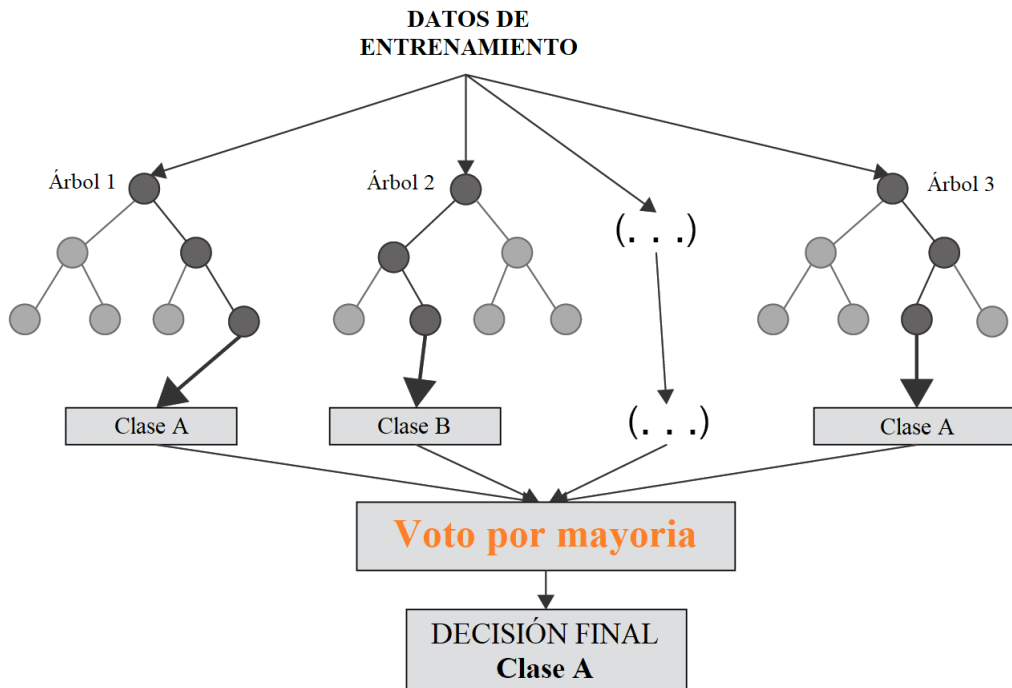


Figura 5.5: Esquema del funcionamiento interno de un algoritmo *Random Forest* [51].

Para conseguir esta diversificación de los modelos entre sí y la no correlación entre los árboles, el algoritmo RF aplica los siguientes métodos [52]:

- **Bagging:** Cualquier pequeño cambio en los datos de entrenamiento puede dar lugar a un resultado significativamente distinto. Es modelo se aprovecha de esto con la asignación de datos de forma aleatoria y con reemplazo a cada árbol, dando lugar a resultados diferentes y, por tanto, árboles no correlacionados.
- **Aleatoriedad de las características:** En el proceso de división de los nodos de un árbol de decisión normal se toman todas las características y se selecciona la que consiga una separación más grande entre los nodos izquierdo y derecho. Sin embargo, en cada árbol individual del bosque aleatorio solo es posible seleccionar características de un subconjunto aleatorio. De esta forma, la mutación entre los distintos árboles está aún más forzada consiguiendo una mayor diferenciación entre ellos.

La implementación del algoritmo se ha llevado a cabo a través de Scikit-Learn, una librería perteneciente a Python de ML. Es una herramienta de uso muy sencillo y de código abierto lo cual permite que la puesta en práctica del algoritmo sea fácil y simple.

Otro aspecto a considerar es el sobreajuste de la red, también conocido como *overfitting*. La creación de un modelo de ML no solo consiste en suministrar multitudes de datos. Este hecho puede converger en defectos que posteriormente perjudiquen la precisión o el rendimiento del algoritmo. Un modelo está sobreajustado cuando al entrenarlo introducimos muchos datos, de forma que, en vez de categorizarlos correctamente, comienza a aprenderlos de memoria debido al alto grado de detalle. Como consecuencia, cuando al modelo le introduzcamos un conjunto de datos distinto al que se ha aprendido, la predicción pasará de ser perfecta a fallar.

El algoritmo RF colabora positivamente impidiendo el sobreajuste gracias a la aleatoriedad mencionada anteriormente. A pesar de ello, se ha utilizado la validación cruzada o *Cross Validation* para alcanzar una alta generalización y asegurar la independencia de los resultados obtenidos. Generalmente, métodos como éste son empleados en modelos que no cuentan con altas cantidades de datos en los subconjuntos de Train y Test, como el caso de este proyecto.

El número de métodos existente para llevar a cabo la validación cruzada es abundante, si bien en este proyecto se ha utilizado *Repeated Stratified K-Fold*, otra herramienta disponible en Scikit-Learn[53]. El procedimiento que lleva a cabo este método consiste en hacer K divisiones del conjunto de datos en partes iguales. El modelo utiliza el primer subconjunto en la primera iteración para aplicar la validación cruzada y los K-1 restantes para entrenar. En la segunda iteración se cambia el subconjunto que se emplea para validar y el resto de nuevo se utilizan en el entrenamiento del modelo. Este proceso se repite hasta que todas las subdivisiones son utilizadas durante la validación. Finalmente, los resultados obtenidos en cada una de las K repeticiones se promedian obteniendo una única evaluación del modelo ya generalizada.

En la Figura 5.6 se visualiza un esquema del funcionamiento del algoritmo. El conjunto de datos está dividido en 10 partes iguales. En azul está representado el subconjunto utilizado para validar y los otros nueve para entrenar. La estimación de la precisión del modelo se calcula como un promedio de los valores E de cada una de las iteraciones aplicando la siguiente fórmula:

$$E = \frac{1}{K} \sum_{i=1}^K E_i \quad (5.2)$$

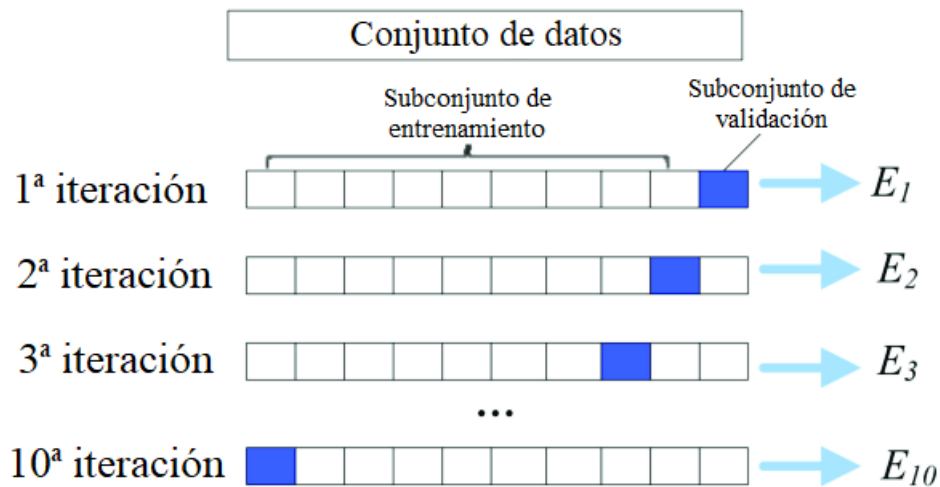


Figura 5.6: Diagrama de validación cruzada para $K = 10$ [54].

5.4.1 Optimización de los Hiperparámetros

Para corregir y maximizar la precisión y el buen rendimiento del modelo utilizado, podemos llevar a cabo una optimización de los hiperparámetros de éste. Estos hiperparámetros consisten en configuraciones del modelo RF que serán diferentes para cada conjunto de datos, es por eso que buscaremos en este paso los mejores para ajustar nuestro modelo [55]. Existe una larga lista de parámetros que se podrían ajustar. En nuestro caso, serán los siguientes:

- `n_estimators`: El número de árboles que va a haber en el RF.
- `max_features`: El número de características de cada partición.
- `max_depth`: El número máximo de particiones que el árbol puede tener.
- `min_samples_split`: El número mínimo de muestras requeridas para que un nodo pueda dividirse.
- `min_samples_leaf`: El número mínimo de muestras requeridas en cada nodo hoja al final de cada uno de los árboles.
- `bootstrap`: Si queremos o no que las muestras se repartan aleatoriamente y con reemplazamiento a la hora de clasificar.

En primer lugar, se ha utilizado el módulo RandomSearchCV perteneciente a Scikit-Learn [56], que tiene en cuenta rangos amplios de valores para cada uno de los hiperparámetros. En la Tabla 5.1 podemos ver la relación entre los hiperparámetros explicados anteriormente y el rango de valores sobre el que se realiza la búsqueda aleatoria.

| Hiperparámetro | Rango de valores |
|-------------------|------------------|
| n_estimators | {200 - 1000} |
| max_features | {log2, sqrt} |
| max_depth | {1 - 15} |
| min_samples_split | {2 - 50} |
| min_samples_leaf | {2 - 50} |
| bootstrap | {True, False} |

Cuadro 5.1: Rango de valores por hiperparámetro.

Existen innumerables combinaciones posibles de estos hiperparámetros. Para cada iteración, se llevará a cabo el mismo método de validación cruzada explicado anteriormente, con 5 subconjuntos y 10 repeticiones. Esto hace un total de 50 realizaciones.

La selección de los hiperparámetros se lleva a cabo atendiendo al mejor comportamiento de todas las realizaciones llevadas a cabo. En la Tabla 5.2 podemos ver cuáles han sido las 10 mejores combinaciones de hiperparámetros ejecutadas en el algoritmo de búsqueda aleatoria y su desempeño en el entrenamiento por medio de la validación cruzada.

| Realización | n_estimators | min_samples_split | min_samples_leaf | max_features | max_depth | bootstrap | Score medio |
|-------------|--------------|-------------------|------------------|--------------|-----------|-----------|-------------|
| 1 | 314 | 50 | 23 | log2 | 7 | False | 0.790519 |
| 2 | 257 | 23 | 2 | log2 | 6 | True | 0.789688 |
| 3 | 542 | 34 | 7 | sqrt | 13 | False | 0.783329 |
| 4 | 428 | 39 | 18 | sqrt | 12 | True | 0.781170 |
| 5 | 942 | 50 | 50 | sqrt | 12 | False | 0.597802 |
| 6 | 828 | 44 | 28 | log2 | 13 | True | 0.597802 |
| 7 | 771 | 12 | 28 | log2 | 2 | False | 0.597802 |
| 8 | 314 | 28 | 34 | sqrt | 3 | True | 0.597802 |
| 9 | 485 | 50 | 39 | log2 | 9 | True | 0.597802 |
| 10 | 828 | 28 | 50 | sqrt | 3 | True | 0.597802 |

Cuadro 5.2: 10 mejores realizaciones llevadas a cabo en la búsqueda aleatoria.

Sin embargo, este tipo de optimización puede no ofrecernos los mejores resultados ya que los valores se comprueban de forma aleatoria. Con el fin de conseguir los valores más apropiados, se han creado gráficos de barras representando en el eje de abscisas los valores explorados de cada hiperparámetro y en el eje de ordenadas la puntuación media de los modelos realizados en cada uno de los valores (Ver Figura 5.7). En la Tabla 5.3 podemos ver la selección elegida de los valores que mejor score medio ofrecen.

| Hiperparámetro | Valores elegidos |
|-------------------|------------------|
| n_estimators | {257, 428, 542} |
| max_features | {log2, sqrt} |
| max_depth | {6, 7} |
| min_samples_split | {23, 34, 39} |
| min_samples_leaf | {2, 7, 18, 23} |
| bootstrap | {False, True} |

Cuadro 5.3: Valores elegidos de cada hiperparámetro.

Tras utilizar RandomSearchCV, podemos implementar otro módulo de Scikit-Learn llamado GridSearchCV [57]. La diferencia con respecto al anterior es que este lleva a cabo una combinación individual de todos ellos siendo el coste computacional mucho mayor. Los valores que se introducen en esta nueva cuadrícula son aquellos que mejores prestaciones han tenido en el algoritmo RandomSearchCV, es decir, los valores indicados en la Tabla 5.3. De esta forma, se buscarán los mejores hiperparámetros de los que ya se consideraron los mejores, optimizando aún más el resultado final. En la Tabla 5.4 vemos cuáles han sido los valores elegidos finalmente para cada uno de los hiperparámetros explorados.

| Hiperparámetro | Valor óptimo |
|-------------------|--------------|
| n_estimators | 542 |
| max_features | sqrt |
| max_depth | 6 |
| min_samples_split | 23 |
| min_samples_leaf | 23 |
| bootstrap | False |

Cuadro 5.4: Valores óptimos introducidos en el Random Forest de cada hiperparámetro.

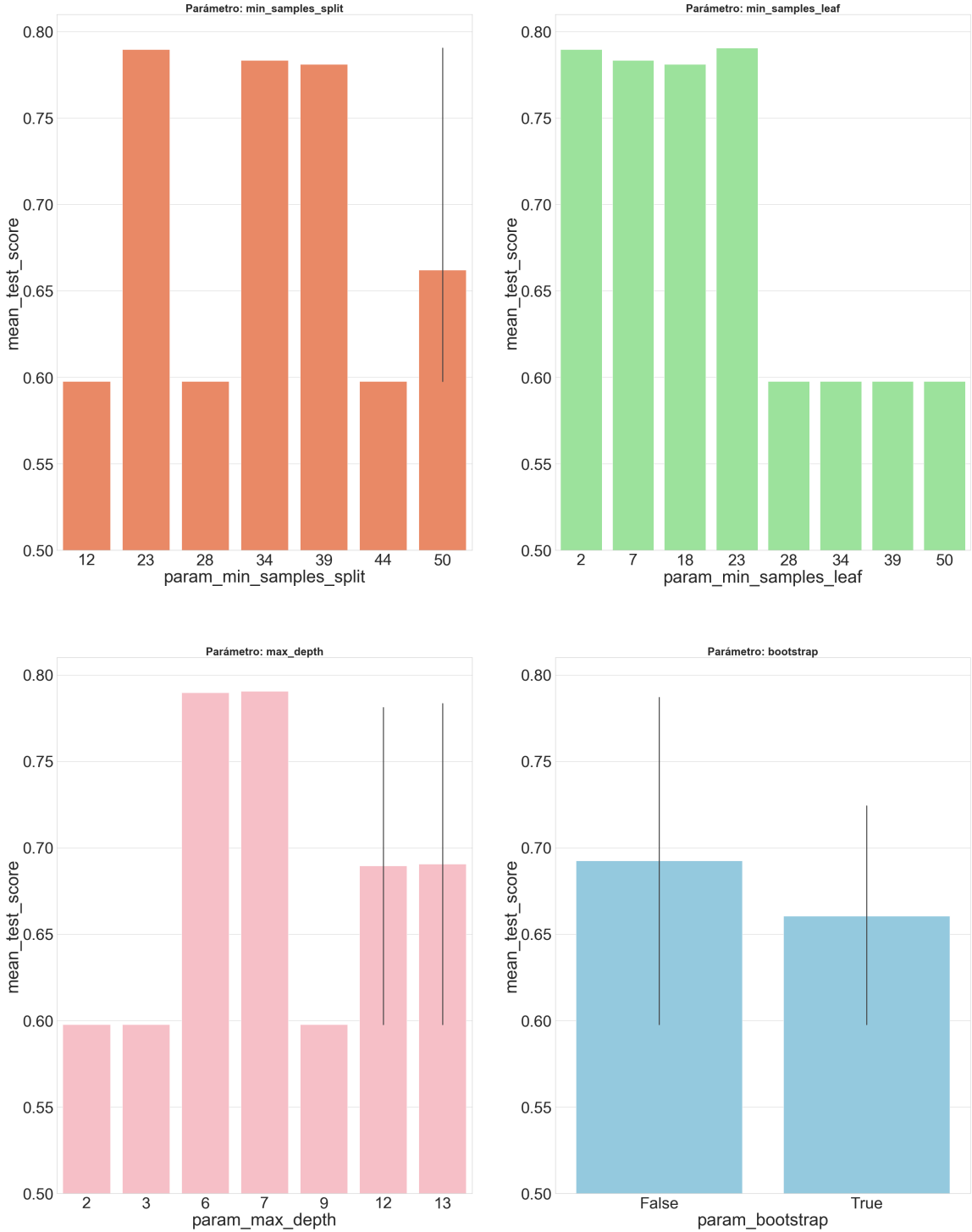


Figura 5.7: Ejemplo de los mejores valores de los hiperparámetros.

5.5 Evaluación del Modelo

Para evaluar el desempeño del clasificador se utiliza la matriz de confusión, una herramienta que permite calcular figuras de mérito que muestran los aciertos o errores que el modelo está adquiriendo durante el aprendizaje [58]. En cada fila de la matriz se disponen los valores de predicción y en cada columna los valores reales, de forma que obtenemos cuatro posibles resultados:

- Verdadero positivo (VP): El modelo predice como positivo un valor que realmente es positivo.
- Falso positivo (FP): El modelo predice como positivo un valor que realmente es negativo.
- Verdadero negativo (VN): El modelo predice como negativo un valor que realmente es negativo.
- Falso negativo (FN): El modelo predice como negativo un valor que realmente es positivo.

El modelo estará acertando cuando detecte VP y VN y estará fallando en los otros casos. Por tanto, cuanto menor sea la proporción de estos últimos con respecto de los primeros, mejores serán los resultados.

A partir de estos cuatro valores surgen la mayoría de las figuras de mérito utilizadas en este proyecto [59]:

- Accuracy o Exactitud: Es el porcentaje casos que el modelo ha clasificado correctamente.

$$\frac{VP + VN}{VP + VN + FP + FN} \quad (5.3)$$

- Precision o Precisión: Es el ratio de verdaderos positivos sobre el total de positivos que el clasificador predice. Con ella podemos medir la calidad del modelo.

$$\frac{VP}{VP + FP} \quad (5.4)$$

- **Recall o Sensibilidad:** Es la capacidad del algoritmo de clasificación de detectar respuesta al tratamiento en los pacientes que realmente la han tenido. Esta métrica también es conocida como Tasa de Verdaderos Positivos.

$$\frac{VP}{VP + FN} \quad (5.5)$$

- **Specificity o Especificidad:** Es la capacidad del clasificador de detectar la proporción de casos negativos que realmente son negativos. En este caso, a esta métrica también se le conoce como tasa de Verdaderos Negativos.

$$\frac{VN}{VN + FP} \quad (5.6)$$

- **F1_score:** Combina la precisión y la sensibilidad en un único valor.

$$2 \times \frac{Precision \cdot Recall}{Precision + Recall} \quad (5.7)$$

- **ROC/AUC:** Ambas son métricas que evalúan el rendimiento del clasificador [60].

1. **ROC:** Se trata de una curva de probabilidad definida por la sensibilidad. Indica la capacidad del modelo de distinguir entre un paciente con respuesta al tratamiento y un paciente que no ha respondido.

$$ROC = 1 - Especificidad = \frac{FP}{VN + FP} \quad (5.8)$$

2. **AUC:** Es el área que existe bajo la curva ROC. Cuanto mayor sea su valor, mayor potencial tendrá el clasificador de diferenciar las dos clases de pacientes.

En la Figura 5.8 se puede ver una clasificación ideal donde la línea verde representa a todos los pacientes sin respuesta al tratamiento (VN) y la roja representa a aquellos que sí han respondido parcial o totalmente al tratamiento (VP). Ambas curvas no se superponen por lo que el modelo posee una medida ideal de separación, es decir, acertará en el 100% de los casos siendo perfectamente capaz de distinguir entre las dos clases sin equivocarse. Como consecuencia, el valor de la sensibilidad siempre será 1 resultando en una AUC = 1.

Sin embargo, en la Figura 5.9 se visualiza un ejemplo cuyas curvas se superponen introduciendo de esta forma errores de predicción, es decir, los falsos positivos y negativos. Todos aquellos

5.5. EVALUACIÓN DEL MODELO

valores correspondientes a la curva verde que estén por encima del umbral establecido serán falsos positivos ya que realmente son negativos pero el modelo los clasifica como positivos. De la misma manera, aquellos pertenecientes a la curva roja que caigan por debajo del umbral serán los falsos negativos, puesto que el modelo los predice como negativos siendo en la realidad positivos. Dependiendo del posicionamiento del umbral, vamos a conseguir optimizar estos errores de predicción. En el caso de esta imagen, el valor del AUC es 0.7, lo que significa que existe un 70% de probabilidades de que el modelo sea capaz de distinguir entre pacientes con respuesta al tratamiento y pacientes sin respuesta, es decir, positivos y negativos.

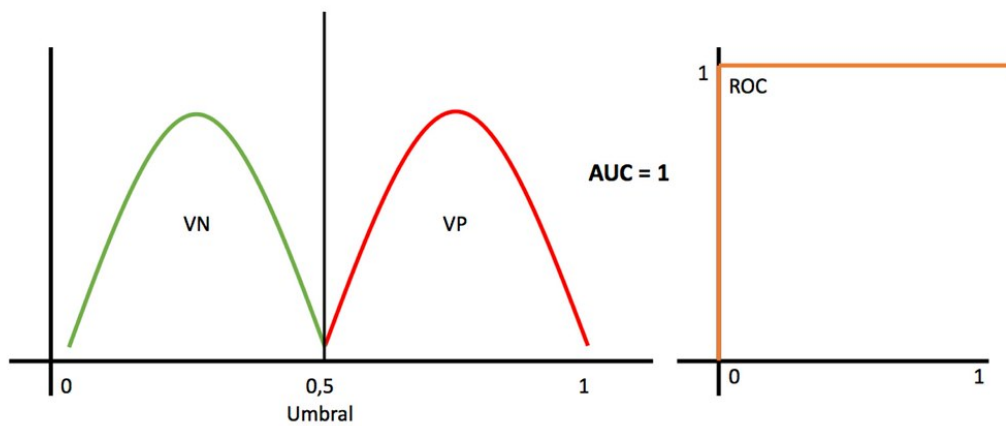


Figura 5.8: Curva ROC-AUC en una situación ideal de clasificación [60].

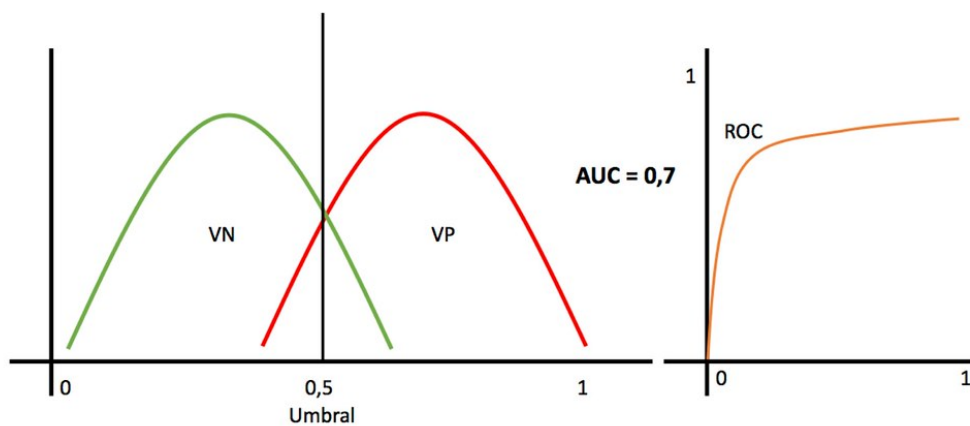


Figura 5.9: Curva ROC-AUC en una situación real de clasificación [60].

5.6 Explicabilidad del Modelo

Generalmente, un modelo de ML elaborado puede generar predicciones precisas, pero éstos ciertamente tienen una esencia natural de caja negra. La interpretabilidad de un algoritmo de aprendizaje automático es necesaria para contribuir a su adopción. La herramienta SHAP (SHapley Additive exPlanations) resulta de gran utilidad para incrementar la transparencia de cualquier algoritmo [61].

Para llevar esto a cabo, se utilizará *TreeExplainer*, una clase incluida dentro del módulo SHAP indicada para explicar algoritmos ensemble como es el caso del RF empleado en este proyecto. En la Figura 5.10 se muestra el funcionamiento del algoritmo SHAP, donde este toma como entradas el conjunto de datos y los valores obtenidos en la predicción y establece una explicación basada en los valores de Shapley.

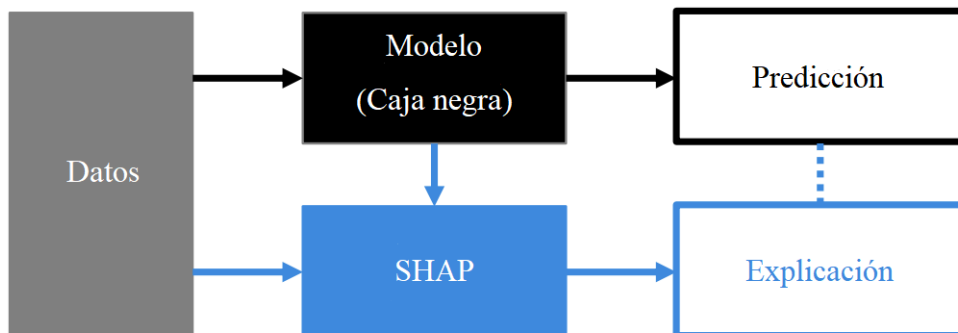


Figura 5.10: Esquema del flujo de trabajo de SHAP [61].

Capítulo 6

Resultados

A continuación, se presentan los hallazgos obtenidos en esta investigación, tanto del clasificador empleado como de las características más relevantes.

6.1 Rendimiento del Clasificador

Los resultados de las métricas calculadas para evaluar el modelo se muestran en la Tabla 6.1. Para todos ellos se ha computado la media de los valores obtenidos en la validación cruzada.

| Métrica | Valor |
|-------------|-----------------|
| Recall | 0.75 ± 0.12 |
| Specificity | 0.60 ± 0.20 |
| Accuracy | 0.64 ± 0.07 |
| Precision | 0.71 ± 0.08 |
| F1-Score | 0.71 ± 0.09 |
| AUC | 0.70 ± 0.12 |

Cuadro 6.1: Resultados de la clasificación empleada.

Además, en la Figura 6.1 se observa la curva ROC obtenida con el rendimiento del modelo que indica la capacidad de este clasificar pacientes con respuesta al tratamiento y sin respuesta al tratamiento. En la misma vemos una recta discontinua roja que representa un modelo hipotético sin capacidad de discriminación entre ambas clases, cuyo valor de AUC es de 0.5 ya que divide el área total en dos partes. La línea azul muestra el valor de AUC del modelo entrenado con el clasificador Random Forest. A medida que la curva aumente su área, mayor será la capacidad discriminativa de clasificador.

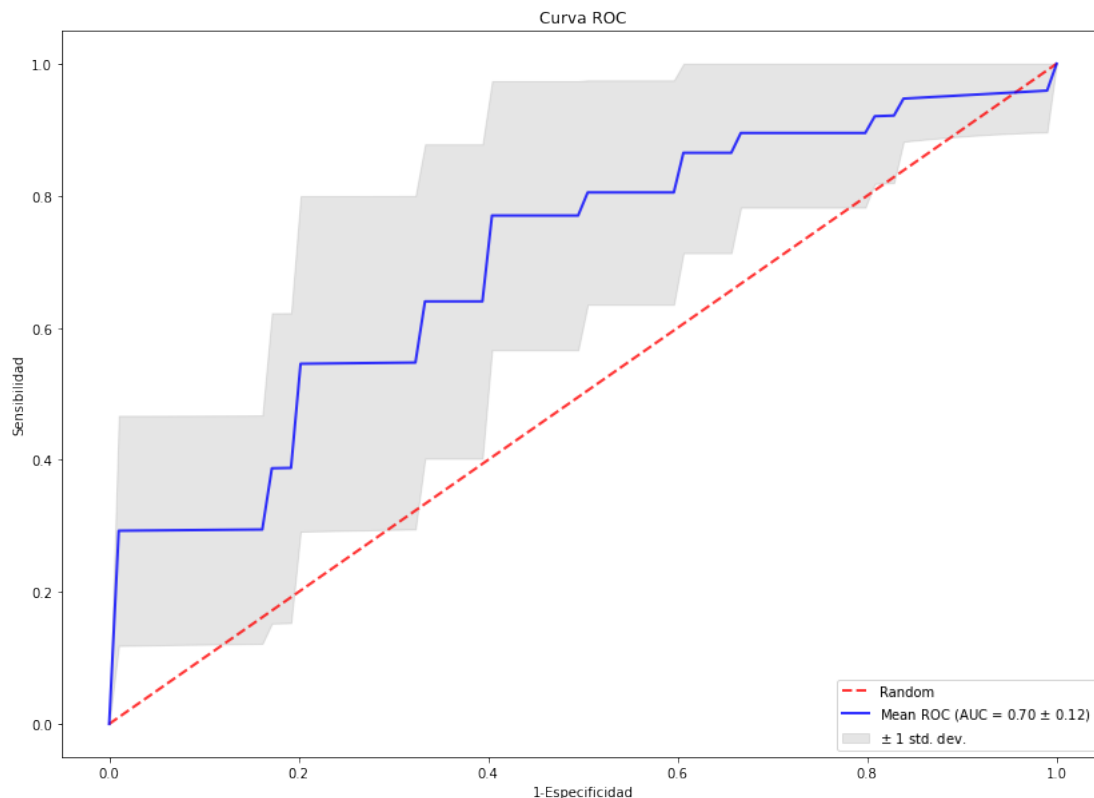


Figura 6.1: Curva ROC obtenida tras el entrenamiento del Random Forest.

6.2 Estudio de las Características

Puesto que contamos con un número muy alto de características radiómicas, es importante llevar a cabo un estudio de éstas. Así, se va a valorar la importancia de cada una de ellas en el proceso de discriminación entre pacientes con respuesta y sin respuesta al tratamiento empleado. Tal y como se explicó en la sección 5.6, se ha utilizado la herramienta SHAP para visualizar cuáles son las que más información aportan al estudio.

Como observamos en la Figura 6.2, hay 5 variables que superan el valor de 0.015, consideradas las más importantes por el algoritmo SHAP. Tan solo una de éstas es de la imagen original y, si nos fijamos en el resto de variables del gráfico, todas ellas son derivadas, lo que nos sugiere un mejor funcionamiento del algoritmo gracias a los filtros aplicados.

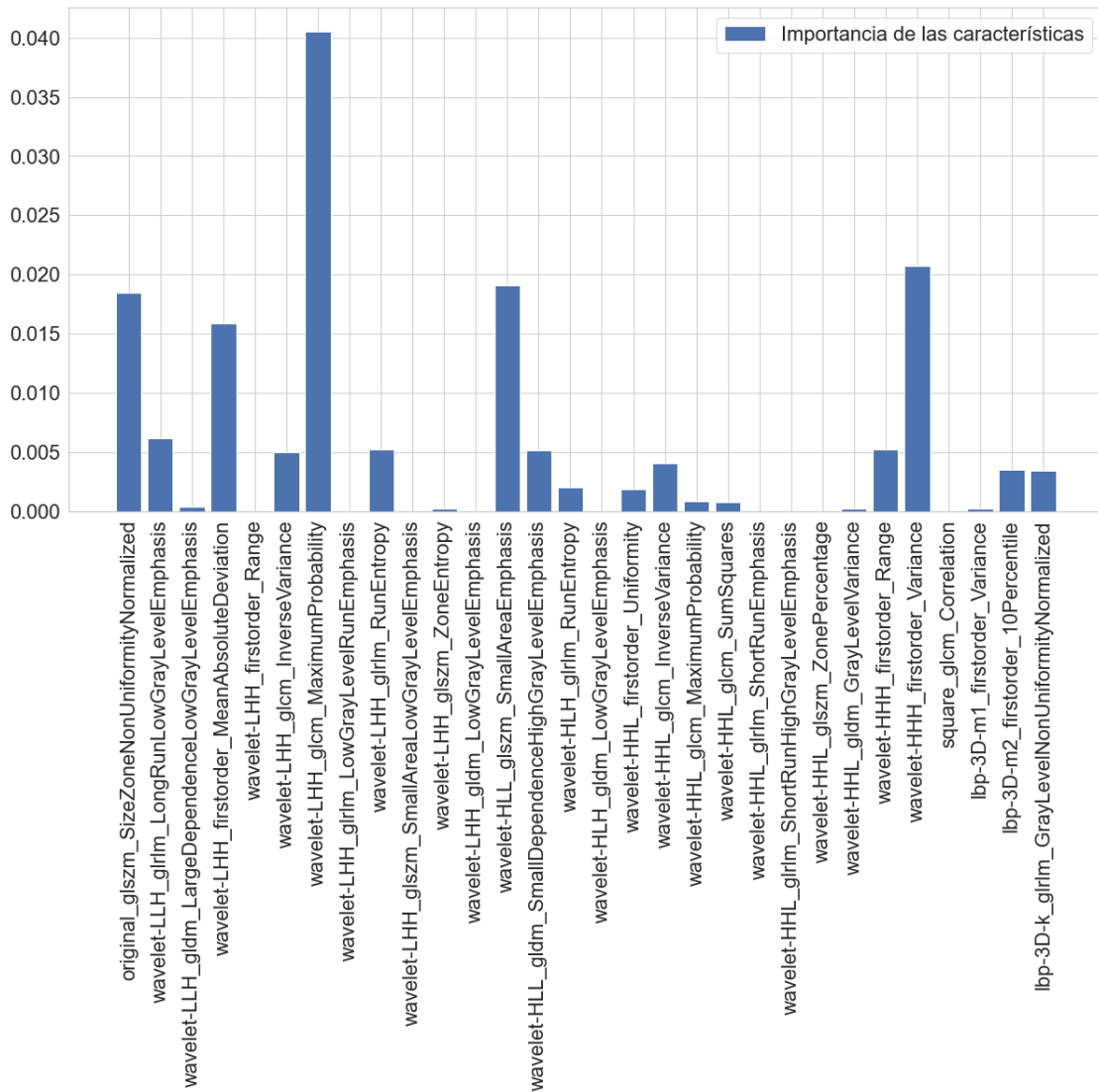


Figura 6.2: Gráfico con la descripción global de las características calculada como la suma de todos los valores absolutos de SHAP para cada una de las variables.

En la Figura 6.3 podemos visualizar la separabilidad entre las dos clases de pacientes para cada una de las características por medio de histogramas, de tal manera que obtengamos una visión de la distribución de los valores de las características en función de la clase a la que pertenezca cada paciente.

6.2. ESTUDIO DE LAS CARACTERÍSTICAS

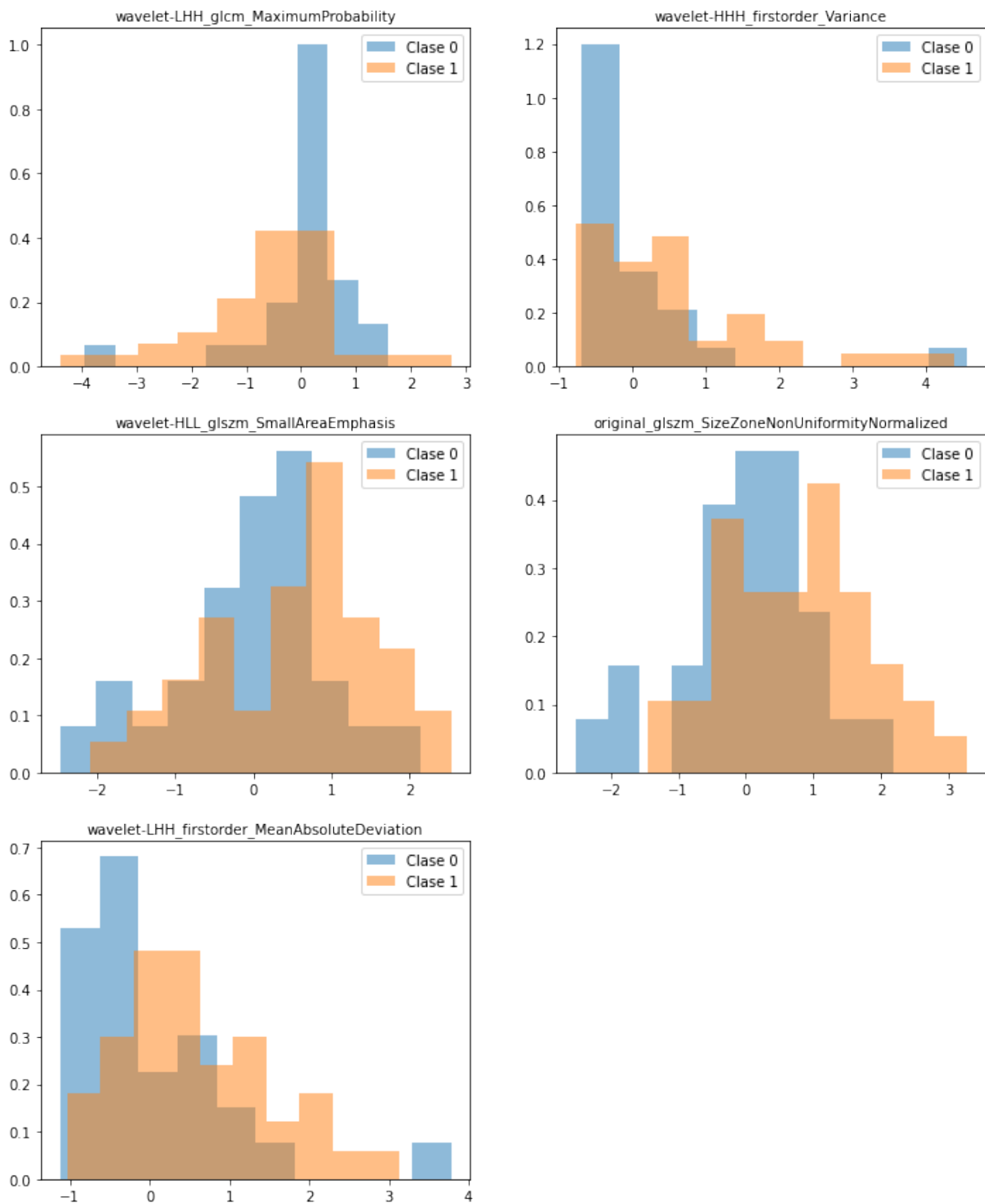


Figura 6.3: Histogramas de las cuatro características más importantes según el algoritmo SHAP.

Capítulo 7

Discusión

El cáncer de pulmón de células no pequeñas es una de las enfermedades oncológicas con mayor mortalidad e incidencia, pese a los avances tecnológicos para su diagnóstico y posterior tratamiento. Por ello, resulta de vital importancia la detección temprana de todos los nódulos. Existen múltiples modalidades de imágenes médicas que pueden aplicarse durante el proceso de detección. No obstante, la tomografía computarizada se ha convertido en la herramienta principal para la localización de nódulos pulmonares. Posteriormente, la labor del radiólogo consiste en determinar todos los tumores malignos que existan analizando sus dimensiones, forma o localización. Este procedimiento es enormemente complejo y requiere el máximo nivel de atención y tiempo. Buscando una solución a dicha problemática surge la radiómica.

Bajo este punto de vista, este campo novedoso trata de extraer datos y llevar a cabo un análisis profundo y avanzado de las imágenes médicas por medio de algoritmos matemáticos. Dicho concepto subyace de la idea de que el ojo humano es incapaz de percibir cierto tipo de datos cuantitativos que pueden esconder información sobre la fisiopatología del tejido. Uno de los cometidos más importantes de la radiómica es la elaboración de modelos de aprendizaje automático capaces de predecir la respuesta al tratamiento de un paciente, basándose en las características radiómicas extraídas de imágenes de TAC. Este nuevo planteamiento es clave para conseguir una medicina adaptada, donde se personalice el tratamiento de acuerdo con las características individuales de cada paciente y de sus tumores.

La respuesta radiológica es la columna vertebral de la respuesta al tratamiento en pacientes con CPCNP. En la práctica, su evaluación consiste fundamentalmente en el análisis del tamaño tumoral, aunque adicionalmente se apoya en otro tipo de características del tumor como su homogeneidad. Cuantitativamente, este punto de vista pasa por alto gran cantidad de información de la imagen. La radiómica cuenta con el poder de detectar indicadores cuantitativos de la respuesta al tratamiento en las imágenes médicas con el fin de ajustar al paciente el tratamiento con anterioridad en el transcurso de la enfermedad para perfeccionar los resultados. En este

proyecto se ha explorado la capacidad de la radiómica combinada con modelos de inteligencia artificial para predecir la respuesta de la quimio-radioterapia en pacientes con CPCNP. Además, se han estudiado dichas características y su importancia para conseguir explicar los resultados obtenidos.

7.1 Impacto de la Segmentación

Todas las características son extraídas de los volúmenes segmentados, por lo que la segmentación manual de los tumores es determinante en el proceso radiómico. En algunos tipos tumorales, este paso resulta una problemática debido a que muchos nódulos poseen bordes difuminados y, por tanto, difíciles de delimitar. Este método supone un alto nivel de conocimiento médico e implica máxima precisión y, en ocasiones, gran cantidad de tiempo de trabajo.

Por otro lado, en un proceso de segmentación manual, como es el caso de este proyecto, el operador es quien acota las regiones de interés bajo su propio criterio. Esto consume gran cantidad de tiempo y, a causa de la dependencia de la persona encargada de la segmentación, se introduce un alto nivel de variabilidad inter-sujeto en los resultados obtenidos. Es decir, distintos operadores podrán llevar a cabo una segmentación de la misma imagen de distintas formas, consiguiendo unos resultados diferentes entre una y otra. Además, resulta bastante complicado que un mismo operador segmente de la misma manera una imagen dos veces, incluyendo así aún más variabilidad, en este caso intra-sujeto.

Se han impulsado nuevas metodologías de segmentación automática y semiautomática por medio de distintas regiones anatómicas. Las condiciones previas integran los conceptos mencionados anteriormente: rentabilidad del tiempo, precisión y reproducibilidad de los bordes. Algunos de estos métodos utilizan algoritmos de crecimiento de regiones, donde el usuario debe definir un punto dentro de la región tumoral de interés que se quiere segmentar [62]. Como se puede intuir, se trata de un método basado en la similitud de píxeles. El problema reside en aquellas lesiones no homogéneas en las que el operador debe hacer una corrección vigorosa.

Se han desarrollado otros métodos de segmentación más robustos basados en la estadística. STAPLE (Simultaneous Truth and Performance Level Estimation) es un ejemplo de estos [63]. Este algoritmo tiene en cuenta un repertorio de muchas segmentaciones de una sola imagen y, con esta información, estima estadísticamente la segmentación real acompañada de un cálculo

del rendimiento que se obtiene en cada una de las segmentaciones. La fuente de cada una de las segmentaciones iniciales que se introducen en el algoritmo pueden ser provenientes de segmentaciones automáticas o bien manuales llevadas a cabo por varios operadores. En un estudio llevado a cabo por Warfield SK [64] se ha demostrado la capacidad real que tiene este método y sus buenos resultados en la clínica.

Una gran variedad de algoritmos diferentes llevan a cabo esta tarea pero aún no se ha determinado el método universal para aplicaciones de este tipo que implican el uso de imágenes médicas, y se continúa investigando otros nuevos que resuelvan las limitaciones. De hecho, unas características pueden mostrar gran reproducibilidad e invariabilidad con el uso de un método, pero no con otro.

7.2 Rendimiento del Clasificador

El sistema elegido para clasificar entre respuesta, o no, al tratamiento se basa en el algoritmo RF. Se han obtenido los valores de las métricas calculadas que se indican a continuación: *recall* (0.75), *specificity* (0.60), *accuracy* (0.64), *precision* (0.71), f1-score (0.71) y AUC (0.70). A la vista de estos resultados es razonable pensar que el enfoque radiómico expone una modesta capacidad para la predicción del resultado de los tratamientos y la respuesta a estos en pacientes con CPCNP. Cuando se trata de aquellos sujetos que no lo harán, se comete un porcentaje de error en la predicción aún mayor. Es por ello que todavía no conforma una herramienta lo suficientemente valiosa para su implementación clínica.

La asociación de respuesta, o no, a cada paciente podría conformar una posible explicación a estos resultados. Por un lado, los pacientes han sido etiquetados tras la visualización por parte del médico de las imágenes del TAC post-tratamiento, incluyendo en gran medida un componente de subjetividad: el sesgo humano. No se han llevado a cabo pruebas complementarias del tumor, ni nuevas intervenciones para comprobar la veracidad del etiquetado. Por otro lado, inicialmente en este estudio se hizo una división de los pacientes en cuatro grupos dependiendo de su respuesta al tratamiento:

- Respuesta total: Desaparición completa del tumor.
- Respuesta parcial: Reducción de parte del tumor.

- Estabilidad: No se visualiza ni mejora ni empeoramiento del tumor.
- Progreso: Aumento del tumor.

Los dos primeros se agruparon en una misma clase: respuesta al tratamiento, etiquetándolos con un 1. Los dos últimos, por el contrario, se agruparon en otra clase: No respuesta al tratamiento, etiquetándolos con un 0. Esta unificación de pacientes conlleva una mayor variabilidad dentro de cada clase, donde un tumor con respuesta parcial puede asemejarse en mayor medida a un paciente con estabilidad que a un paciente cuyo tumor ha respondido totalmente. Además, si hacemos un recuento de la cantidad de pacientes existentes en cada clase, tan solo 1 de los 40 respondedores ha obtenido una respuesta total, por lo que se hace la clasificación aún más imprecisa.

En cuanto al RF, un estudio llevado a cabo por Parmar *et al.* [65] muestra que el modelo escogido para la clasificación contribuye en gran medida a la variación en el rendimiento de éste. En su investigación se pusieron a prueba 12 modelos diferentes de aprendizaje automático (redes neuronales, RF, máquinas de vectores soporte, etc.), experimentando con datos radiómicos. Los investigadores concluyeron que los algoritmos basados en RF suponen la mejor elección para gestionar la inestabilidad de las características extraídas.

A pesar del éxito y la relevancia de los métodos *ensemble* en el campo de la radiómica, la evolución paralela de algoritmos basados en DL constituyen un núcleo de investigación sólido. Avanzo *et al.* [35] señala que el DL es capaz de ofrecer una gran mejora en la segmentación y detección de tumores pulmonares y en la clasificación de los mismos. No hay duda de que en un futuro próximo se alcanzará una combinación exitosa de la radiómica con el ML en el ámbito del CP, que dotará de una herramienta de ayuda al diagnóstico potente y precisa.

Los métodos basados en IA, como el DL nombrado anteriormente, son superiores al planteamiento radiómico habitual [66]. Esto se debe a que con la aplicación de la IA se descarta la selección manual de características imposibilitando los sesgos humanos. En cambio, la IA implica la necesidad de obtener conjuntos de datos considerablemente grandes, donde la racionalidad escondida tras las decisiones tomadas resultan una caja negra.

Atendiendo al tipo de imágenes, la TAC es una de las tecnologías para diagnóstico con imágenes más ampliamente utilizadas en la detección del CP actualmente. Su combinación con el PET ha demostrado ser la mejor opción en la estadificación de la enfermedad debido a su alto rendi-

miento [67]. La literatura existente sobre la aplicación de la radiómica para estimar la respuesta terapéutica del paciente y predecir los resultados se fundamenta en imágenes anteriores al tratamiento. A este respecto, la TAC proporciona características anatómicas que si las integramos con la información de los exámenes PET, obtenemos una completa cantidad de datos también metabólicos lo suficientemente útiles para identificar signos de recurrencia o detectar la enfermedad precozmente. Un estudio sugiere que las características de los exámenes PET cuentan con una precisión mayor y una capacidad de predicción más alta que las características radiómicas basadas en TAC [68]. Además, aunque estas últimas tienen una modesta capacidad para distinguir la respuesta tumoral, si que mejoran frente al ojo humano. Por lo tanto, las imágenes multimodales constituirían una herramienta idónea para predecir la recurrencia tumoral.

7.3 Estudio de las Características

Con el paso del tiempo, las características radiómicas se han vuelto sustancialmente un instrumento muy apreciado en cuanto a la detección de tumores y su evolución. Su notoria popularidad se pone de manifiesto por el número exponencialmente creciente de nuevas publicaciones que incorporan este término. Sin embargo, a pesar de su potencial, todavía presenta ciertos desafíos que resolver antes de su traducción a la práctica clínica. Si bien, es de esperar que la radiómica influya paulatinamente en la terapia de los tumores pulmonares, optimizando la cadena de diagnóstico-tratamiento-seguimiento.

En una revisión publicada en 2021 se resume la situación del ámbito de la radiómica en la actualidad analizando el resultado de 14 estudios en los que se indaga en los predictores radiómicos de la respuesta al tratamiento de CPCNP [69]. En esta publicación, se concluye que existe una gran variedad de indicadores, desde características basadas en el histograma como la simetría, hasta características derivadas relacionadas con la textura como la media del filtro Wavelet. Rara vez se ha identificado una característica como predictor común de respuesta al tratamiento en todos los estudios evaluados. Esto explica la inestabilidad que existe con las características radiómicas. En el futuro, será decisivo establecer la reproducibilidad y el beneficio clínico de los modelos radiómicos de forma que se demuestre su calidad como mecanismo de toma de decisiones para facilitar el tratamiento personalizado de pacientes con CPCNP.

En este proyecto se ha dado gran relevancia al análisis y explicabilidad de los resultados en base

a las características más importantes en el proceso de clasificación. De las 1562 características que obtuvimos con el extractor de Py-Radiomics, se eliminaron las constantes y todas aquellas con una correlación muy alta conservando 30 de ellas que fueron las que finalmente participaron en el proceso de clasificación. Son tan solo las primeras de estas las que reúnen un gran valor explicativo para el clasificador y, por tanto, las más relevantes en el problema que se aborda.

En concreto, cuatro de ellas están relacionadas con el filtro Wavelet. La característica restante es extraída de la imagen original y pertenece al grupo GLSZM. Por tanto, vamos a comentar el significado de cada una de las cinco por separado:

- **wavelet-LHH_glcm_MaximumProbability:** Consiste en un tipo de imagen Wavelet en la que se ha aplicado un filtro paso bajo al eje X y un filtro paso alto a los ejes Y, Z. Esta característica evalúa la aparición del par de valores de intensidad más predominante en píxeles vecinos.
- **wavelet-HHH_firstorder_Variance:** En este caso se aplica un filtro Wavelet con filtro paso alto en los tres ejes. La varianza calcula la distancia del nivel de intensidad de cada vóxel con el valor medio de intensidad de toda la imagen.
- **wavelet-HLL_glszm_SmallAreaEmphasis:** De nuevo, consiste en un subtipo de las imágenes con filtro Wavelet. Sin embargo, el filtro paso alto se ha aplicado en el eje X y dos filtros pasos bajos en los ejes restantes. Esta característica mide la distribución de zonas pequeñas y texturas más finas.
- **original_glszm_SizeZoneNonUniformityNormalized:** Esta es la única característica original que se ha considerado como importante desde el inicio. Evalúa la diferencia de volúmenes que existe en la imagen completa en su versión normalizada. Cuanto menor sea este valor, más homogénea es la imagen.
- **wavelet-LHH_firstorder_MeanAbsoluteDeviation:** Esta característica es del mismo subtipo que la primera mencionada. En este caso, se mide la distancia media que existe, en valor absoluto, desde el valor de intensidad de cada uno de los píxeles hasta el valor medio de la imagen completa.

De acuerdo con estos resultados, resulta de gran valor analizar el tipo de características que mejor explican el estudio. Como hemos visto anteriormente, dos de ellas son de primer orden [70], basadas en el histograma de niveles de gris global, es decir, se estudia un único vóxel, en vez de

todos en conjunto. Otras dos características importantes pertenecen a la clase GLSZM [71], que aporta información sobre la distribución espacial de píxeles consecutivos con el mismo nivel de gris, formando la base de una matriz con estos. En consecuencia, las imágenes con texturas más homogéneas, conformarán matrices más anchas y planas que las texturas heterogéneas.

Por último, nos encontramos con que la característica más importante, con gran diferencia, pertenece a la clase de características GLCM [72]. Esta matriz describe la función de probabilidad conjunta en la región de la imagen delimitada por la máscara, es decir, el tumor. En este caso, no se basa tanto en propiedades físicas de la imagen sino en valores estadísticos, concretamente para esta característica, la probabilidad máxima en intensidades vecinas.

Esta exploración de las características aporta resultados destacables. En primer lugar, y atendiendo al tipo de características que se han presentado como más importantes, cabe destacar que son las imágenes filtradas las que más valor radiómico han aportado en el proceso de clasificación. En consecuencia, se puede deducir que el radiólogo al disponer solo de la imagen original, puede estar perdiendo información indispensable. En segundo lugar, se pone en valor la utilidad de la radiómica ya que si esto no fuera así, no se podría diferenciar a simple vista aquellas características que han aportado mayor beneficio. Por otro lado, se debe poner de manifiesto el tipo de tarea que se está llevando a cabo. El conjunto de datos contiene imágenes tumorales todos ellos malignos, esto implica que muchas de las características radiómicas serán compartidas entre el grupo de pacientes que responden al tratamiento y los que no lo hacen, tan solo por ser ambos tumores patológicos y, con seguridad, es posible que los resultados bajos en algunas de las métricas sean consecuencia de esto.

Capítulo 8

Conclusión y Trabajos Futuros

La utilización de la radiómica como herramienta de ayuda al diagnóstico ha causado poco a poco una revolución de métodos innovadores de evaluación de enfermedades difíciles de llevar a cabo si no fuera por la ayuda de la computación. Cada vez se están abarcando más ámbitos dentro de la medicina en los que aplicar todos estos conocimientos, sobre todo en la esfera de la radiología. Lo más valioso de todo esto es que el objetivo es prestar ayuda a los profesionales sanitarios expertos en la materia, nunca reemplazarlos. Cada día los resultados están mejorando gracias a los avances tecnológicos que culminarán con una precisión próxima al 99% en un futuro no muy lejano.

8.1 Conclusión

Como ya se se anunció en los Objetivos, con el presente proyecto se ha pretendido diseñar una herramienta de clasificación para predecir la respuesta al tratamiento de los pacientes con cáncer de pulmón de células no pequeñas. Y no solo eso, sino que se ha aportado explicabilidad al algoritmo que a simple vista resulta una caja negra de gran rendimiento, consiguiendo entender las características de mayor valor para el modelo de clasificación.

La primera parte del trabajo pretende unificar en pocas páginas el fundamento médico y teórico del CP, que puede mejorar la interpretabilidad y la pertinencia de este proyecto agrupando ambas visiones: la medicina y el aprendizaje automático.

Por otro lado, el presente trabajo proporciona no solo las herramientas y técnicas necesarias para implementar los conocimientos adquiridos, sino que se aporta valor añadido permitiendo que el lector interesado en continuar un proceso deductivo en su aprendizaje, sea capaz de conseguir una base de análisis y abstracción de la radiómica y el aprendizaje automático, dotándole de

autonomía en la lectura de textos más avanzados.

Los métodos utilizados han proporcionado unos resultados humildes pero que han servido para extraer cuáles han sido las limitaciones de este proyecto. Bajo este punto de vista, es imprescindible destacar el tamaño del conjunto de datos. Por un lado, sería de gran utilidad incrementar el tamaño muestral, debido al número tan elevado de características radiómicas que existen, muchas de ellas superfluas. En cuanto a esto último, encontrar un método óptimo de selección de las características menos correlacionadas entre sí aumentaría en gran medida la capacidad predictiva del modelo utilizado. Añadido a esto, existe un problema de capacidad de cómputo que hace imposible la extracción de todas las características originales y derivadas de ciertos pacientes cuya imagen ocupa un volumen muy grande de datos, disminuyendo aún más la muestra de pacientes disponible. Por esta razón, la búsqueda de las mejores características ha sido exhaustiva durante todo el desarrollo del proyecto, de modo que se evite al máximo el sobreentrenamiento.

8.2 Trabajos Futuros

La radiómica tiene un futuro largo y alentador como instrumento de apoyo clínico. La manera en la que se ha desarrollado este trabajo permite llevar a cabo un seguimiento detallado de la metodología para que cualquier persona ajena interesada en profundizar en esta investigación sea capaz de seguir todos los pasos llevados a cabo y mejorarla.

Futuras líneas de trabajo que se puedan construir sobre esta investigación podrían dirigirse a la implementación de algoritmos predictivos no supervisados, es decir, modelos capaces de aprender sin el control humano. Así, se conseguirá una nueva forma de tomar decisiones gracias a la cantidad masiva de datos de entrada, dejando el trabajo del especialista para otras labores que requieran mayor especialización. Por otro lado, guarda una estrecha relación con los resultados el método de segmentación elegido, una vía por la que profundizar para conseguir un proceso más sencillo, rápido y automático.

Realmente son muchas las posibilidades que se abren como fruto de la experiencia adquirida durante la evolución de este trabajo. El código se encuentra documentado para un mejor entendimiento y un enfoque metodológico ampliamente detallado a lo largo de todo este. El enlace es:

https://github.com/saraltuna/sara_altuna_tfg_ing_biomedica

Que mi humilde aportación ayude a incentivar el interés en este campo de la misma manera que los meses dedicados han provocado en mi.

Breve nota sobre la dedicación de la autora

Durante el inicio de mi etapa estudiando el grado en Ingeniería Biomédica estuve siempre muy perdida. Yo quería ser médico. Poco a poco fue creciendo mi interés en el ámbito de la Inteligencia Artificial y el análisis de las imágenes médicas, y como una máquina podía ser capaz de llevar a cabo tareas tan sumamente trascendentales en la vida de las personas. Precisamente ese era el único objetivo que siempre tuve claro que quería perseguir: dar ayuda a los demás.

A mitad del camino empecé a ser consciente de que quería seguir aprendiendo sobre este campo y, pese a que en ese momento de la carrera no hubiese explotado aun todos esos conceptos, yo tenía claro que quería empaparme de experiencia y nuevos conocimientos. Porque este grado te permite justamente eso, una base que poder utilizar más tarde para elegir el camino que más te apasione.

Fue entonces cuando tuve la suerte de conocer a Ángel. No tardé mucho en darme cuenta de la inspiración tan grande que sus clases estaban suponiendo para mí. Sin que él ni yo lo supiéramos, me guió hasta que entendí hacia donde quería dirigirme. Entonces decidí escoger este tema para mi último trabajo.

Ya todo parecía relativamente fácil. Ahora solo tenía que sacar el máximo provecho a todo lo que aprendí durante la carrera y tratar de preparar algo que aportase valor a un campo en el que siempre estuve interesada: la medicina.

Han sido unos meses de gran aprendizaje donde he crecido tanto personal como profesionalmente. Este trabajo no ha hecho más que aumentar mi entusiasmo por la ingeniería en combinación con la medicina. Ahora me siento preparada para comenzar la siguiente etapa.

Aprovecho estas últimas líneas para agradecer a todo compañero y profesor que he tenido la suerte de conocer en el camino y de los que pude aprender algo nuevo. Y gracias por haber dedicado tu tiempo a estas páginas.

Bibliografía

- [1] A. d. Córdoba 15. Cáncer de Pulmón España. Consultado el 1 de junio de 2023. Disponible en: <http://cancer-pulmon.com/>.
- [2] Anatomía de los pulmones y del aparato respiratorio | Cigna. Consultado el 1 de junio de 2023. Disponible en: <https://www.cigna.com/es-us/knowledge-center/hw/anatoma-de-los-pulmones-y-del-aparato-respiratorio-zm2676>.
- [3] Cáncer de pulmón. Consultado el 1 de junio de 2023. Disponible en: <https://medlineplus.gov/spanish/lungcancer.html>.
- [4] Cáncer de pulmón de células no pequeñas - Introducción. Consultado el 1 de junio de 2023. Disponible en: <https://www.cancer.net/es/tipos-de-c%C3%A1ncer/c%C3%A1ncer-de-pulm%C3%B3n-de-c%C3%A9lulas-no-peque%C3%B1as/introducci%C3%B3n>.
- [5] IARC – International Agency for Research on Cancer. Consultado el 1 de junio de 2023. Disponible en: <https://www.iarc.who.int/>.
- [6] SEOM: Sociedad Española de Oncología Médica - SEOM. Consultado el 1 de junio de 2023. Disponible en: <https://seom.org/>.
- [7] Home | Redecan. Consultado el 1 de junio de 2023. Disponible en: <https://redecana.org/es>.
- [8] Observatorio de resultados - Portal de Salud de la Comunidad de Madrid. Consultado el 1 de junio de 2023. Disponible en: <http://observatorioresultados.sanidadmadrid.org/GraficosEstadoPoblacion.aspx?ID=11>.
- [9] D. M. Pérez, D. R. Ysamat, D. L. Muñoz, D. J. Algar, D. J. L. Romeo, and D. I. Barneto, “Anatomía Patológica Radiología Neumología Cirugía Torácica Oncología Radioterápica Oncología Médica.”
- [10] Riesgos para la salud al fumar tabaco. Consultado el 1 de junio de 2023. Disponible en: <https://www.cancer.org/es/cancer/prevencion-del-riesgo/tabaco/riesgos-para-la-salud-debido-al-tabaquismo/fumar-tabaco.html>.

- [11] Factores de riesgo para el cáncer de pulmón. Consultado el 2 de junio de 2023. Disponible en: <https://www.cancer.org/es/cancer/tipos/cancer-de-pulmon/causas-riesgos-prevencion/factores-de-riesgo.html>.
- [12] F. M. Nirvana *et al.*, “Inherited lung cancer: a review,” *ecancermedicalscience*, vol. 14, 2020.
- [13] Tratamientos. Consultado el 2 de junio de 2023. Disponible en: <https://afectadoscancerdepulmon.com/tratamientos/>.
- [14] La cirugía en el cáncer de pulmón | Roche Pacientes. Consultado el 2 de junio de 2023. Disponible en: <https://rochepacientes.es/cancer/pulmon/tratamiento/cirugia.html>. [Online]. Available: <https://rochepacientes.es/cancer/pulmon/tratamiento/cirugia.html>
- [15] Cirugía de cáncer de pulmón. Consultado el 2 de junio de 2023. Disponible en: <https://www.mayoclinic.org/es/diseases-conditions/lung-cancer/multimedia/lung-cancer-surgery/img-20006167>.
- [16] Radioterapia en el cáncer de pulmón | Roche Pacientes. Consultado el 2 de junio de 2023. Disponible en: <https://rochepacientes.es/cancer/pulmon/tratamiento/radioterapia.html>.
- [17] Nueva técnica revoluciona el tratamiento del cáncer de pulmón. Consultado el 2 de junio de 2023. Disponible en: <https://www.topdoctors.es/articulos-medicos/la-radioterapia-estereotaxica-de-cuerpo-revoluciona-el-cancer-de-pulmon>.
- [18] Quimioterapia en el cáncer de pulmón | Roche Pacientes. Consultado el 2 de junio de 2023. Disponible en: <https://rochepacientes.es/cancer/pulmon/tratamiento/quimioterapia.html>.
- [19] J. Ferreiro, J. García, R. Barceló, and I. Rubio, “Quimioterapia: efectos secundarios,” *Gaceta Médica de Bilbao*, vol. 100, no. 2, pp. 69–74, 2003.
- [20] La inmunoterapia en cáncer de pulmón | Roche Pacientes. Consultado el 2 de junio de 2023. Disponible en: <https://rochepacientes.es/cancer/pulmon/tratamiento/inmunoterapia/inmunoterapia.html>.
- [21] Terapia dirigida – Cáncer de pulmón | Roche Pacientes. Consultado el 2 de junio de 2023. Disponible en: <https://rochepacientes.es/cancer/pulmon/tratamiento/terapia-dirigida.html>.
- [22] Componentes de un Tomógrafo. Consultado el 2 de junio de 2023. Disponible en: <http://www.needgoo.com/componentes-de-un-tomografo/>.

- [23] Tomografía Computarizada (TC). Consultado el 2 de junio de 2023. Disponible en: <https://www.nibib.nih.gov/espanol/temas-cientificos/tomograf%C3%ADa-computarizada-tc>.
- [24] ¿En qué consiste una Tomografía Axial Computarizada (TAC)? - El Tamiz. Consultado el 2 de junio de 2023. Disponible en: <https://eltamiz.com/2008/01/22/%C2%BFen-que-consiste-una-tomografia-axial-computarizada-tac/>.
- [25] M. G. Tume Fuentes, “Estado del arte de la inteligencia artificial y su aplicación en el mantenimiento,” 2022.
- [26] ¿Qué es el aprendizaje supervisado? | IBM. Consultado el 2 de junio de 2023. Disponible en: <https://www.ibm.com/mx-es/topics/supervised-learning>.
- [27] ¿Qué es el aprendizaje no supervisado y cuándo usarlo? Consultado el 2 de junio de 2023. Disponible en: <https://www.inesdi.com/blog/que-es-aprendizaje-no-supervisado/>.
- [28] ¿Qué es Deep Learning? | IBM. Consultado el 2 de junio de 2023. Disponible en: <https://www.ibm.com/es-es/topics/deep-learning>.
- [29] M. E. Mayerhoefer, A. Materka, G. Langs, I. Häggström, P. Szczypiński, P. Gibbs, and G. Cook, “Introduction to radiomics,” *Journal of Nuclear Medicine*, vol. 61, no. 4, pp. 488–495, 2020.
- [30] J. D. Shur, S. J. Doran, S. Kumar, D. Ap Dafydd, K. Downey, J. P. O’Connor, N. Papanikolaou, C. Messiou, D.-M. Koh, and M. R. Orton, “Radiomics in oncology: a practical guide,” *Radiographics*, vol. 41, no. 6, pp. 1717–1732, 2021.
- [31] D. V. Fried, S. L. Tucker, S. Zhou, Z. Liao, O. Mawlawi, G. Ibbott, and L. E. Court, “Prognostic value and reproducibility of pretreatment CT texture features in stage iii non-small cell lung cancer,” *International Journal of Radiation Oncology* Biology* Physics*, vol. 90, no. 4, pp. 834–842, 2014.
- [32] M. Khorrami, P. Prasanna, A. Gupta, P. Patil, P. D. Velu, R. Thawani, G. Corredor, M. Ali-lou, K. Bera, P. Fu *et al.*, “Changes in CT radiomic features associated with lymphocyte distribution predict overall survival and response to immunotherapy in non-small cell lung cancer,” *Cancer immunology research*, vol. 8, no. 1, pp. 108–119, 2020.
- [33] S. U. Berlangieri, A. M. Scott, S. R. Knight, G. J. Fitt, O. F. Hennessy, H. J. Tochon-Danguy, C. Peter Clarke, and W. John McKay, “F-18 fluorodeoxyglucose positron emission tomography in the non-invasive staging of non-small cell lung cancer,” *European journal of cardio-thoracic surgery*, vol. 16, no. Supplement_1, pp. S25–S30, 1999.

- [34] K. Tournoy, S. Maddens, R. Gosselin, G. Van Maele, J. Van Meerbeeck, and A. Kelles, “Integrated FDG-PET/CT does not make invasive staging of the intrathoracic lymph nodes in non-small cell lung cancer redundant: a prospective study,” *Thorax*, vol. 62, no. 8, pp. 696–701, 2007.
- [35] M. Avanzo, J. Stancanello, G. Pirrone, and G. Sartor, “Radiomics and deep learning in lung cancer,” *Strahlentherapie und Onkologie*, vol. 196, pp. 879–887, 2020.
- [36] H. J. Yoon, I. Sohn, J. H. Cho, H. Y. Lee, J.-H. Kim, Y.-L. Choi, H. Kim, G. Lee, K. S. Lee, and J. Kim, “Decoding tumor phenotypes for ALK, ROS1, and RET fusions in lung adenocarcinoma using a radiomics approach,” *Medicine*, vol. 94, no. 41, 2015.
- [37] E. Huynh, T. P. Coroller, V. Narayan, V. Agrawal, Y. Hou, J. Romano, I. Franco, R. H. Mak, and H. J. Aerts, “CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer,” *Radiotherapy and Oncology*, vol. 120, no. 2, pp. 258–266, 2016.
- [38] T. L. Chaunzwa, A. Hosny, Y. Xu, A. Shafer, N. Diao, M. Lanuti, D. C. Christiani, R. H. Mak, and H. J. Aerts, “Deep learning classification of lung cancer histology using CT images,” *Scientific reports*, vol. 11, no. 1, p. 5471, 2021.
- [39] Hospital Ramón y Cajal. Consultado el 3 de junio de 2023. Disponible en: <https://www.comunidad.madrid/hospital/ramonycajal/>.
- [40] Radiodiagnóstico. Consultado el 3 de junio de 2023. Disponible en: <https://www.comunidad.madrid/hospital/ramonycajal/ciudadanos-oferta-asistencial/servicios-centrales/radiodiagnostico>.
- [41] D. Terrón Santos, J. L. Domínguez Álvarez *et al.*, “REGLAMENTO (UE) 2016/ 679 DEL PARLAMENTO EUROPEO y DEL CONSEJO - de 27 de abril de 2016 - relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la directiva 95/ 46/ CE (reglamento general de protección de datos),” 2019.
- [42] 3D Slicer image computing platform. Consultado el 5 de junio de 2023. Disponible en: <https://slicer.org/>.
- [43] Welcome to Python.org. Consultado el 5 de junio de 2023. Disponible en: <https://www.python.org/about/>.

- [44] J. J. Van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, and H. J. Aerts, “Computational radiomics system to decode the radiographic phenotype,” *Cancer research*, vol. 77, no. 21, pp. e104–e107, 2017.
- [45] Project Jupyter. Consultado el 6 de junio de 2023. Disponible en: <https://jupyter.org>.
- [46] Radiomic Features. Consultado el 6 de junio de 2023. Disponible en: <https://pyradiomics.readthedocs.io/en/latest/features.html>.
- [47] K. Pearson, “VII. Note on regression and inheritance in the case of two parents,” *proceedings of the royal society of London*, vol. 58, no. 347-352, pp. 240–242, 1895.
- [48] Correlación de Pearson (formulación). Consultado el 12 de junio de 2023. Disponible en: https://web.archive.org/web/20091215105427/http://www.psico.uniovi.es/Dpto_Psicologia/metodos/tutor.6/fcope.html.
- [49] Selección de características con los scores más altos usando SelectKBest — documentación de cursos de analítica y machine learning -. Consultado el 12 de junio de 2023. Disponible en: <https://jdvelasq.github.io/courses/modulos/sklearn/39%20seleccion%20de%20caracteristicas/notebooks/06%20SelectKBest.html>.
- [50] What is Random Forest? | IBM. Consultado el 12 de junio de 2023. Disponible en: <https://www.ibm.com/topics/random-forest>.
- [51] F. Hernández, *8 Random Forest | Modelos Predictivos*, consultado el 12 de junio de 2023. Disponible en: https://fhernanb.github.io/libro_mod_pred/rand-forests.html.
- [52] T. Yiu. Understanding Random Forest. Consultado el 12 de junio de 2023. Disponible en: <https://towardsdatascience.com/understanding-random-forest-58381e0602d2>.
- [53] RepeatedStratifiedKFold. Consultado el 12 de junio de 2023. Disponible en: https://scikit-learn/stable/modules/generated/sklearn.model_selection.RepeatedStratifiedKFold.html.
- [54] Ten-fold cross validation diagram. Consultado el 12 de junio de 2023. Disponible en: https://www.researchgate.net/figure/Ten-fold-cross-validation-diagram-The-dataset-was-divided-into-ten-parts-and-nine-of_fig1_326465007.

- [55] W. Koehrsen. Hyperparameter tuning the random forest in python. Consultado el 12 de junio de 2023. Disponible en: <https://towardsdatascience.com/hyperparameter-tuning-the-random-forest-in-python-using-scikit-learn-28d2aa77dd74>.
- [56] RandomizedSearchCV. Consultado el 12 de junio de 2023. Disponible en: https://scikit-learn/stable/modules/generated/sklearn.model_selection.RandomizedSearchCV.html.
- [57] GridSearchCV. Consultado el 12 de junio de 2023. Disponible en: https://scikit-learn/stable/modules/generated/sklearn.model_selection.GridSearchCV.html.
- [58] J. I. B. Arce. La matriz de confusión y sus métricas – Inteligencia Artificial –. Consultado el 12 de junio de 2023. Disponible en: <https://www.juanbarrios.com/la-matriz-de-confusion-y-sus-metricas/>.
- [59] Precision, Recall, F1, Accuracy en clasificación - IArtificial.net. Consultado el 12 de junio de 2023. Disponible en: <https://www.iartificial.net/precision-recall-f1-accuracy-en-clasificacion/>.
- [60] L. Gonzalez. Curvas ROC y Área bajo la curva (AUC). Consultado el 12 de junio de 2023. Disponible en: <https://aprendeia.com/curvas-roc-y-area-bajo-la-curva-auc-machine-learning/>.
- [61] Welcome to the SHAP documentation — SHAP latest documentation. Consultado el 12 de junio de 2023. Disponible en: <https://shap.readthedocs.io/en/latest/index.html>.
- [62] M. Wang and D. Li, “An Automatic Segmentation Method for Lung Tumor Based on Improved Region Growing Algorithm,” *Diagnostics*, vol. 12, no. 12, p. 2971, 2022.
- [63] O. Commowick and S. K. Warfield, “Estimation of inferential uncertainty in assessing expert segmentation performance from staple,” *IEEE transactions on medical imaging*, vol. 29, no. 3, pp. 771–780, 2010.
- [64] S. K. Warfield, K. H. Zou, and W. M. Wells, “Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation,” *IEEE transactions on medical imaging*, vol. 23, no. 7, pp. 903–921, 2004.
- [65] C. Parmar, P. Grossmann, J. Bussink, P. Lambin, and H. J. Aerts, “Machine learning methods for quantitative radiomic biomarkers,” *Scientific reports*, vol. 5, no. 1, p. 13087, 2015.

- [66] C. K. Kuhl and D. Truhn, “The long route to standardized radiomics: unraveling the knot from the end,” pp. 339–341, 2020.
- [67] I. Pinilla and N. G. León, “Utilidad de la PET/TC en el cáncer de pulmón,” *Radiología*, vol. 51, no. 3, pp. 248–260, 2009.
- [68] L. Shi, Y. He, Z. Yuan, S. Benedict, R. Valicenti, J. Qiu, and Y. Rong, “Radiomics for response and outcome assessment for non-small cell lung cancer,” *Technology in cancer research & treatment*, vol. 17, p. 1533033818782788, 2018.
- [69] M. R. Chetan and F. V. Gleeson, “Radiomics in predicting treatment response in non-small-cell lung cancer: current status, challenges and future perspectives,” *European radiology*, vol. 31, pp. 1049–1058, 2021.
- [70] Radimics Features First Order. Consultado el 27 de junio de 2023. Disponible en: <https://pyradiomics.readthedocs.io/en/latest/features.html#radiomics.firstorder.RadiomicsFirstOrder>.
- [71] Radimics Features GLSZM. Consultado el 27 de junio de 2023. Disponible en: <https://pyradiomics.readthedocs.io/en/latest/features.html#module-radiomics.glszm>.
- [72] Radimics Features GLCM. Consultado el 27 de junio de 2023. Disponible en: <https://pyradiomics.readthedocs.io/en/latest/features.html#module-radiomics.g lcm>.