Eugeni Bernad Gutiérrez

Desenvolupament de models predictius basats en radiòmica per la resposta als tractaments de radioteràpia estereotàctica extracranial (SBRT) en pacients amb càncer de pulmó

Treball Fi de Grau

dirigit per la Dra. Meritxell Arenas Prat i el Dr. Victor Hernández Masgrau

Grau en Enginyeria Biomèdica

UNIVERSITAT ROVIRA I VIRGILI

Tarragona

2023

Índex

1 Introducció

El càncer de pulmó és un dels més diagnosticats durant el que portem de 2023.

Segons el GECP (Grupo Español de Cáncer de Pulmón) el càncer de pulmó ha augmentat la seva mortalitat [1], doncs els experts alerten d'una pujada del 2.4%, essent igual de preocupant en homes com en dones. Per aquests motius millores en el diagnòstic precoç o a una etapa potencialment curable tindrien un gran impacte a la salut humana.

En el servei de Oncologia Radioteràpica de l'Hospital Sant Joan de Reus treballen amb diferents tècniques amb la fi d'arribar a aquests avenços que ajudin als pacients.

Amb el temps, els radiòlegs han pogut identificar diferents característiques físiques visuals qualitatives que poden diferenciar de lesions benignes d'altres malignes. Alguns exemples podrien ser la mida de la lesió, la atenuació i els espais aeris quístics o perilesionals [2]. Les característiques radiòmiques són descriptors basats en imatges que son capaços de quantificar diferents característiques del tumor. Junt amb la Intel·ligència Artificial (IA), podrem combinar mètodes per a identificar nous biomarcadors que anteriorment no s'han pogut valorar amb la fi d'obtenir un model de predicció o de pronòstic.

1.1 Tractament de radioteràpia estereotàctica extracranial (SBRT)

La SBRT és un tractament que administra la dosis de radiació d'una manera molt precisa i intensa amb l'objectiu de maximitzar les capacitats de la radioteràpia per a destruir el càncer mentre alhora es redueixen els seus efectes secundaris als teixits i òrgans sans [2]. Actualment s'utilitzen tractaments amb SBRT per al tractament de tumors de mida petita i mitjana, malignes o benignes, en llocs molt específics i comuns. Alguns exemples de localitzacions on podem aplicar-ho són, l'abdomen, cap i coll, o el propi pulmó com estarem parlant en aquest document, entre d'altres.

Pel que fa referència a l'administració de la dosis, normalment un tractament de radioteràpia al pulmó, tracta la malaltia amb sessions alternes (dia sí, dia no) de dilluns a divendres durant 15 sessions. Essent un total aproximadament d'uns 37.5 Gy (uns 2.5 Gy per sessió). En el cas de la SBRT les sessions es redueixen a un marge de entre 3 i 5 sessions on la dosis total segueix sent la mateixa. Així llavors estem parlant d'uns 10 Gy per sessió.

Com a qualsevol malaltia que ocupa el servei d'oncologia radioteràpica de l'hospital, en aquest procés participen els tècnics de radioteràpia, que realitzen els TACs de simulació i ajuden al pacient en tot el procés presencial del tractament; l'especialista en oncologia radioteràpica, essent l'especialista mèdic que determina l'àrea a tractar, la dosis que es subministrarà i l'aprovació del tractament, fent continuadament un seguiment del pacient per a valorar el resultat; i un físic, que s'encarregarà de realitzar la dosimetria. Procés on es realitza una simulació del tractament on s'ajustarà la dosis i el funcionament de l'accelerador lineal per tal de "jugar" amb els col·limadors de les màquines i angles de intensitat de dosi per a poder aconseguir irradiar entre el 95-107% del PTV (segmentació que realitza l'especialista en oncologia radioteràpica on s'inclou la malaltia macroscòpica i uns marges per a cobrir la microscòpica) i no superar els màxims establerts per als òrgans de risc i teixits sans.

Com estem parlant d'una tècnica molt precisa, l'administració de la radiació requereix de diferents aparells que ajuden a immobilitzar al pacient de manera que els moviments siguin mínims. Primer, i com a qualsevol tractament de radioteràpia, es realitza un TAC on gràcies a uns tatuatges i junt amb uns làsers de la pròpia sala, es determina la posició que tindrà el

pacient en totes les sessions del tractament. Seguidament utilitzarem un immobilitzador del tòrax que no permetrà que hi hagi molt moviments del tronc superior a causa de la respiració.

Figura 1. Immobilitzador per a la SBRT. Representació de la posició del pacient amb els diferents aparells que ajuden a que es redueixin els moviments naturals de la persona a tractar.

L'accelerador lineal que s'utilitza a l'Hospital Sant Joan de Reus és un accelerador TrueBeam de Varian Medical Systems. Com a tot accelerador lineal produeix radiació d'alta energia (fotons) des de diferents angles del gantry per a maximitzar la dosis a l'àrea tumoral. Aquest, com ens diu la web de Varian, proporciona opcions de tractament flexibles, obtenció de imatges avançades, control de la dosificació d'altra precisió, tractaments optimitzats i una seguretat i fiabilitat alta, fent que el tractament oncològic sigui intuïtiu, innovador i integrat [3].

1.2 Imatge TAC

La tomografia computeritzada és una exploració que produeix múltiples imatges de l'interior del cos que serveix com a examen mèdic de diagnòstic per imatge. Durant l'examen el pacient queda estirat a un llit, que es mou a través d'un anell que conté el equips de raigs X i la computadora. Aquestes imatges es prenen des de diferents angles i s'utilitzen per a crear visualitzacions tridimensionals (3D) dels teixits i òrgans.

Realitzar un TAC proporciona diferents beneficis. Entre ells trobem que ens poden determinar quan és necessària una cirurgia, reduint el nombre de cirurgies exploratòries; i/o millora el diagnòstic i tractament del càncer, gràcies a la ràpida adquisició de les imatges que aporten informació clara i específica [4]. A part es poden obtenir imatges d'una petita part del cos o de tot aquest en un mateix examen, característica que cap altre procediment per imatge combina en una sola sessió.

La computadora realitza una reconstrucció de la imatge gràcies a algorismes. Aquest és un procés de dos passos. Primer, es mesura la transmissió d'un feix de raigs X a totes les trajectòries rectilínies possibles al pla de l'objecte; segon, la atenuació d'un feix de raigs X s'estima en punts de l'objecte. Algorismes que facin aquesta reconstrucció en podem trobar molts, com són el simple backprojection, que és molt ràpid però és molt sensible al soroll; el filtered backprojection, que també és ràpid i té millor desenfocament encara que també pateix amb el soroll i diferents artefactes; series d'expansió, que són relativament lentes però millors envers el soroll; o reconstruccions amb Deep Learning, que son ràpides i bones per a la reducció de dosi [5].

1.2.1 TAC simulador

En el servei de l'Hospital s'utilitza un TAC simulador. Aquest examen es realitza abans de començar el tractament. En certa manera, estem parlant de la part més important, doncs sense aquesta prova no es podria realitzar el estudi de la delimitació de volums i la posició del pacient exacta del dia de tractament. El que es realitzava era una prova TAC sense contrast, on es ficaven una espècie de marcadors situats majoritàriament als laterals de la persona i un

altre cèntric gràcies a l'ajut d'uns làsers que hi havia a la sala. Aquesta marcació venia seguida d'uns tatuatges. Aquests es fan per a que durant la resta de dies de tractaments, el pacient pugui estar col·locat de la mateixa manera, assegurant que la dosi que es deposita va allà on els metges i físics han establert. D'aquí la importància d'aquest primer TAC simulador.

Els tècnics són les persones que decideixen junt amb les indicacions prèvies del metge de consulta, quins artefactes s'utilitzaran per a la posició de cada pacient. Pel que fa a les SBRT, en aquest primer TAC ja s'utilitzen els immobilitzadors de la figura 1, que seran els mateixos que s'usaran els dies de tractament.

1.2.2 TAC de resposta

Un cop finalitzat el tractament, entre 1 mes i 1 mes i mig, es realitza un examen TAC per a que l'especialista valori la resposta al tractament per part del pacient. Aquesta prova ens determinarà si, en el nostre cas, la SBRT ha complert els objectius total, parcialment o per contrari, si el tumor segueix igual de manera estable o inclús amb progressió.

1.3 Radiòmica

La radiòmica és una tècnica de processament d'imatge mèdica que amb l'ajut d'algorismes computacionals passa d'interpretar d'una manera subjectiva a extraure característiques qualitatives i quantitatives d'una secció d'ella, normalment, d'una lesió o tumor detectat [6]. En aquest cas parlem d'imatge en relació al tractament de radioteràpia.

A més a més, [7] la radiòmica ha estat aplicada amb èxit per a predir efectes secundaris com la pneumonitis induïda per la radioteràpia i la immunoteràpia i diferenciar la lesió pulmonar de la recidiva.

Alhora, podem parlar de que la radiòmica com a potenciador a l'hora d'utilitzar i fer més ús del TAC, doncs s'espera que afecti a la pràctica clínica de tractament de tumors pulmonars optimitzant la cadena de diagnòstic-tractament-seguiment de principi a fi.

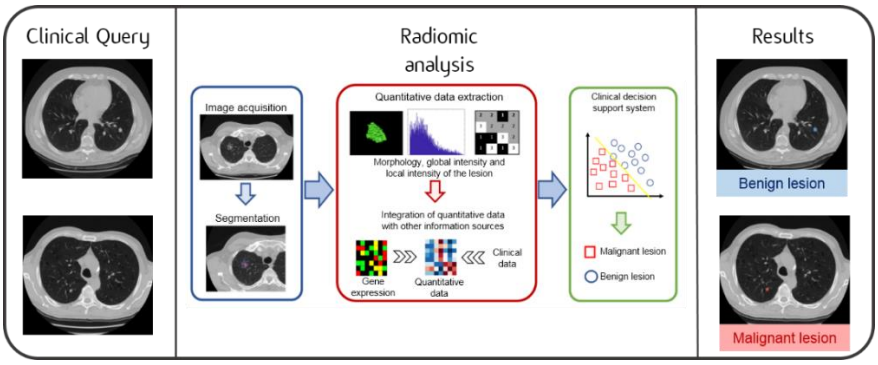


Figura 2. Procés general d'un estudi de radiòmica on diferenciem l'obtenció de la imatge clínica, la quantificació i anàlisi d'aquesta amb el corresponent resultat de predicció.

1.4 Delimitació i obtenció d'imatges

Varian és un a empresa que té molta contribució en el món del càncer. Un dels seus productes de software i sistemes de informació és ARIA.

ARIA és un sistema d'informació per a la oncologia radioteràpica mèdica i quirúrgica. Entre les seves característiques podem trobar que completa la gestió de la informació del pacient, revisant imatges clíniques, receptes, proves de laboratori, entre d'altres; connectivitat fluida, oferint compatibilitat amb HL7 i DICOM; ràpid accés a la informació; obtenció d'imatges

integrada, ajudant optimitzar les tècniques de tractament; i assistència amb el compliment normatiu.

2 Objectius

L'any 2021 es va fer un TFG similar que va permetre l'extracció de característiques radiòmiques a partir de les imatges de TAC d'una subsèrie d'aquests pacients i el desenvolupament d'uns primers models predictius. L'objectiu del la pràctica actual és refinar l'extracció de característiques radiòmiques, ampliar l'estudi a la col·lecció completa de pacients tractats, millorar els models predictius obtinguts i avaluar les seves aplicacions pràctiques potencials per millorar l'atenció clínica dels pacients amb càncer de pulmó, un de els càncers amb més incidència actualment.

3 Materials i mètodes

3.1.1 Smart Segmentation

La segmentació de la nòdul pulmonar a analitzar és essencial fer-la el més precisa possible, ja que una petita variació del volum es pot traduir en desajustos en els valors de les nostres característiques radiòmiques. Al costat de l'ajuda de Víctor Hernández i Mauricio Múrcia, hem fet una segmentació intel·ligent del volum gràcies al propi programari de l'ARIA de Varian. Aquesta segmentació majoritàriament inclou el volum del tumor observat al TAC de simulació previ al tractament de radioteràpia.

Les lesions més properes a les costelles han estat les que més dificultats han causat, doncs el programa detectava la intensitat de l'ós com si fos del nòdul. Això ha provocat que de manera manual hàgim de ajustar la segmentació fins a obtenir únicament el tumor.

Aquesta segmentació és un volum una mica més ajustat que el GTV (Gross Tumor Value). En alguns dels casos el volum que hem fet servir per a l'anàlisi sí que és el corresponent al seu GTV encara que en la majoria de casos utilitzem aquest nou volum.

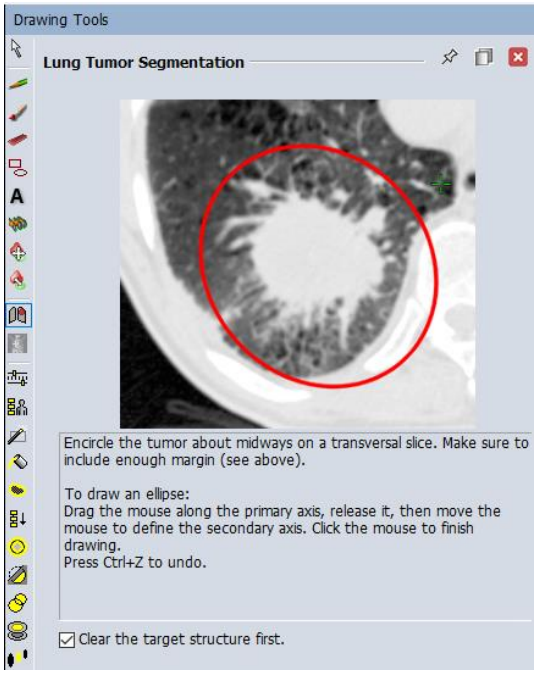


Figura 3. Smart segmentation tool de l'ARIA de Varian Medical Systems (Mètode d'ús).

Primerament obrim l'eina de l'Smart Segmentation d'ARIA de Varian Medical Systems on trobarem una breu explicació del seu funcionament

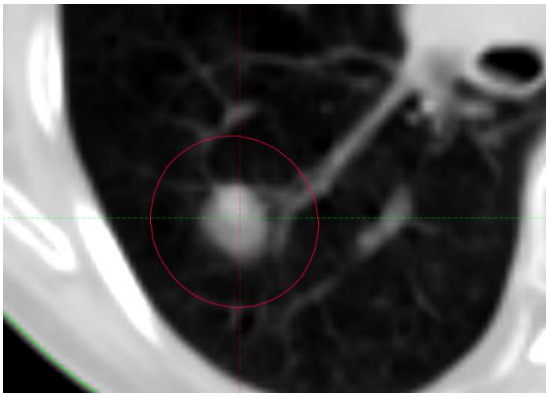


Figura 4. Smart segmentation tool de l'ARIA de Varian Medical Systems.

Tot de seguit obrim l'eina de segmentació, envoltem amb un tall transversal la zona tumoral assegurant-nos d'incloure un marge suficient. Per fer-ho, arrossegarem el 'mouse' al llarg de l'eix primari del tumor fent un click per a acabar de dibuixar. Seguidament, definirem l'eix secundari de manera transversal al primer.

Automàticament el programa començarà a delimitar el tumor que tenim dins de l'el·lipse dibuixada anteriorment arribant finalment a la delineació de l'estructura.

Figura 6. Smart segmentation tool de l'ARIA de Varian Medical Systems.

Per a una millor pràctica, farem zoom d'aquesta delimitació per a corregir possibles errors de la Smart Segmentation, doncs es pot donar el cas de que delimiti zones no tumorals, per exemple, les costelles; o que hagi part del tumor que l'eina no hagi considerat maligna encara que ho sigui.

Finalment, guardem la nova estructura com a 'Smart_pre', i l'especialista mèdic la revisarà per tal de confirmar que la delimitació s'ha realitzat correctament.

3.1.2 Imatge DICOM

Les imatges del TAC es poden extreure del sistema hospitalari en format DICOM (Digital Imaging and Commmunications in Medicine). Aquest és un estàndard internacional per a les imatges mèdiques que permet compartir les imatges mèdiques entre diferents sistemes

d'atenció mèdica sense perdre qualitat ni integritat de la imatge. Per a cada lesió corresponia un TAC amb una segmentació 'Smart_pre' en format DICOM.

Les imatges DICOM poden tenir informació mèdica sensible sobre la salut d'un pacient, doncs és molt fonamental que amb l'objectiu de mantenir la privacitat i la confidencialitat dels pacients, s'anonimitzin les imatges.

Aquestes les exportem a un USB des del sistema informàtic del servei de manera anonimitzada per a poder analitzar-les la nostra plataforma.

3.2 3D Slicer image computing platform

3D Slicer és un programari gratuït de codi obert per a la visualització, el processament, la segmentació, el registre i l'anàlisi d'imatges i malles mèdiques, biomèdiques i d'altres tipus en 3D, així com per a la planificació i la navegació de procediments guiats per imatges.

La versió d'Slicer 3D utilitzada és la Slicer 5.2.1. Per a la visualització hem necessitat instal·lar 2 extensions del programa.

SlicerRT és un conjunt d'eines de radioteràpia per a 3D Slicer, que conté funcions genèriques de RT per a importació/exportació, anàlisi, visualització, amb l'objectiu de fer de 3D Slicer una potent plataforma de recerca en radioteràpia. El desenvolupament de SlicerRT està finançat actualment per CANÀRIE.

SlicerRT es va crear originalment amb finançament de Cancer Care Ontario i el Consorci d'Ontario per a Intervencions Adaptatives en Oncologia Radioteràpica (OCAIRO) per proporcionar un conjunt d'eines gratuïtes i de codi obert per a radioteràpia i intervencions relacionades guiades per imatges. [8]

SlicerRadiomics és mòdul d'extensió per a 3D Slicer que encapsula la biblioteca de Pyradiòmics, que alhora implementa el càlcul d'una varietat de característiques radiòmiques [9].

Amb elles dues més els mòduls inclosos en el propi programa a l'instalar-ho, ja podíem visualitzar en 3D el TAC amb totes les delimitacions de les estructures i extreure les característiques radiòmiques.

3.2.1 Visualització de la Smart Segmentation

En primer lloc, hem importat les imatges DICOM extretes de la base de dades de l'Hospital Sant Joan de Reus a 3D Slicer utilitzant el mòdul 'Add DICOM Data'.

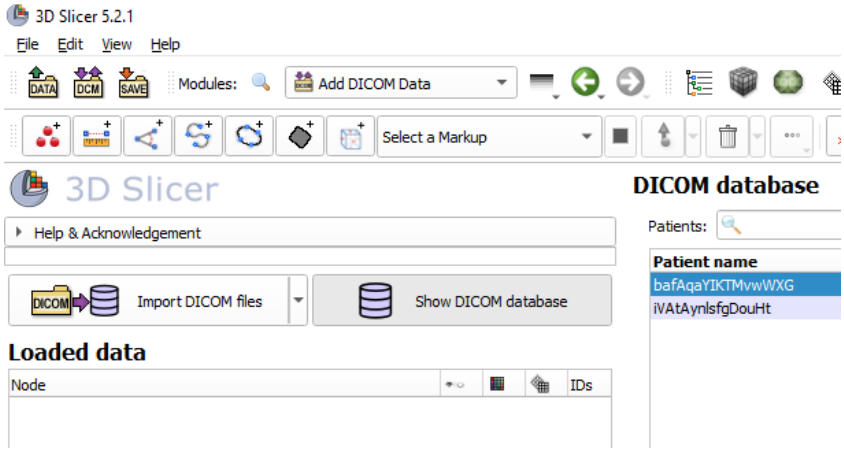


Figura 8. Plataforma 3D Slicer.

En tots els casos i per a preservar la seguretat de les dades, els arxius dels pacients estaven anonimitzat.

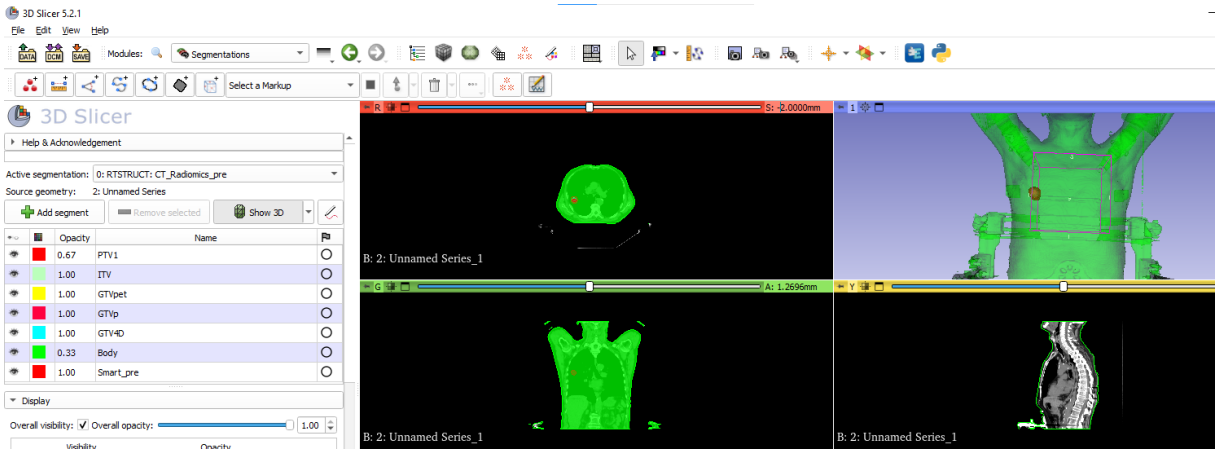


Figura 9. Plataforma 3D Slicer.

Després, utilitzant el mòdul anomenat 'Segmentations' tindrem la primera visualització del TAC i totes les delimitacions que es representen. El següent pas és eliminar totes les estructures no desitjades fins a que només quedi la 'Smart_pre', que és la que hem creat nosaltres per a l'estudi.

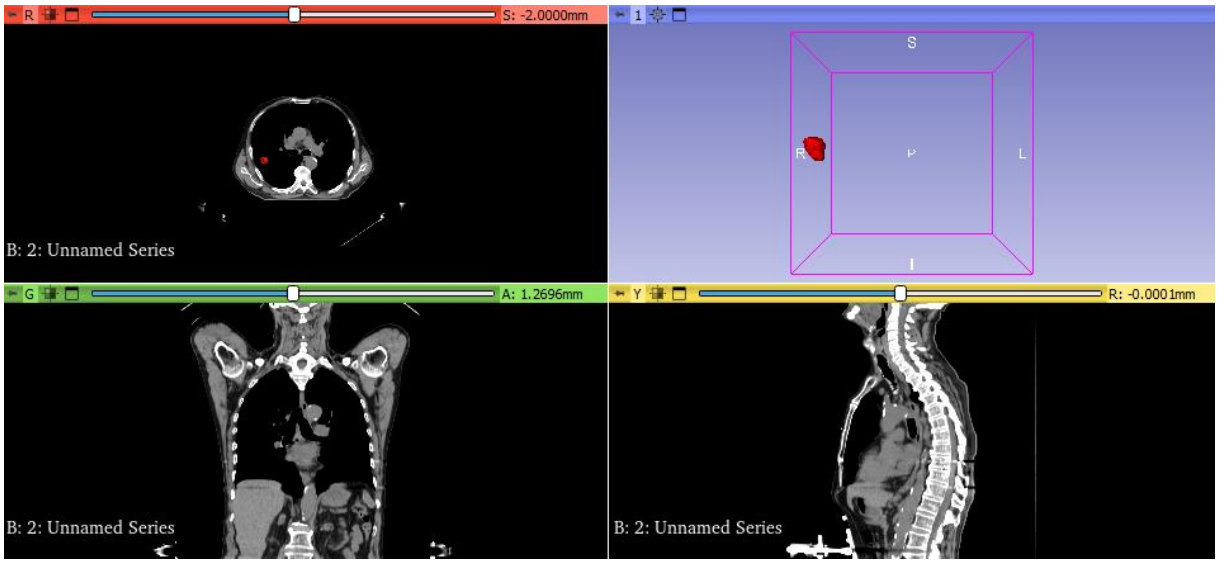


Figura 10. Plataforma 3D Slicer.

A l'instant veurem com les estructures eliminades ja no es visualitzen i només queda el referent a la segmentació intel·ligent.

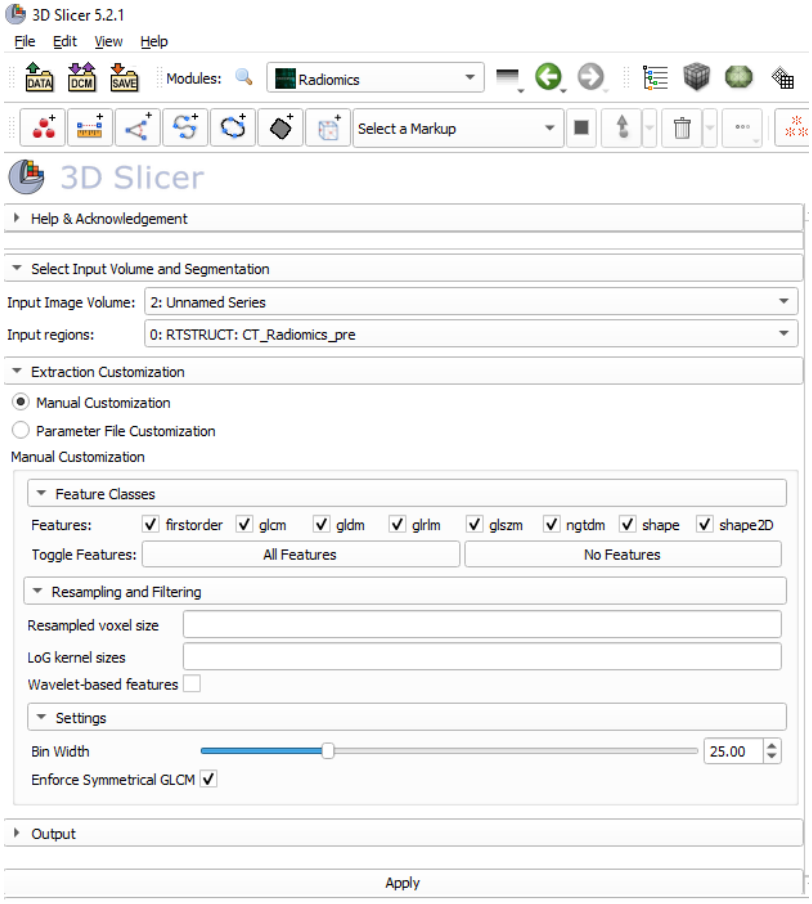


Figura 11. Plataforma 3D Slicer.

Seguidament al mòdul 'Informatic -> Radiomics' escollirem la regió creada anteriorment a Input Regions i escollirem a Manual Customization totes les classes de característiques possibles. Finalment farem click a Apply per a fer la extracció.

Tot aquest procés es farà de manera repetida per a tots els pacients de l'estudi.

3.3 Característiques radiòmiques

Les característiques radiòmiques són descriptors estadístics que caracteritzen l'aspecte visual macroscòpic de la imatge. Les característiques que extraiem són les següents: [10]

Classe	Nombre de característiques
First order	18
GLCM	24
GLDM	14
GLSZM	16
GLRLM	16
NGTDM	5
shape	14

Taula 1. Es visualitzen per consola una quantitat de 107 característiques radiòmiques. Aquestes es poden dividir en diferents classes: Firstorder, gray level co-ocurrence matrix (GLCM), gray level dependence matrix (GLDM), gray level run size zone matrix (GLSZM), gray level run lenght matrix (GLRLM), neighbouring gray tone difference matrix features (NGTDM), shape.

a) Característiques First Order

Són característiques que descriuen la distribució de les intensitats dels vòxels dins de la lesió [10]. Les característiques que formen part d'aquest grup són les següents:

- 10Percentile
- 90Percentile
- Energy
- Entropy
- InterquartileRange
- Kurtosis
- Maximum
- MeanAbsoluteDeviation
- Mean
- Median
- Minimum
- Range
- RobustMeanAbsoluteDeviation
- RootMeanSquared
- Skewness
- TotalEnergy
- Uniformity
- Variance

b) Característiques gray level co-ocurrence matrix (GLCM)

La matriu de co-ocurrencia descriu la freqüència d'aparició d'un parell de valors de l'escala de grisos en una orientació determinada i mesuren la textura d'un tumor [10]:

- Autocorrelation
- ClusterProminence
- ClusterShade
- ClusterTendency
- Contrast
- Correlation
- DifferenceAverage
- DifferenceEntropy
- DifferenceVariance
- Id
- Idm
- Idmn
- Idn
- Imc1
- Imc2
- InverseVariance
- JointAverage
- JointEnergy
- JointEntropy
- MCC
- MaximumProbability
- SumAverage
- SumEntropy
- SumSquares

c) Característiques gray level dependence matrix (GLDM)

Quantifica les dependències del nivell de gris en una imatge, és a dir, el número de vòxels connectats que depenen d'un vòxel central [10]:

- DependenceEntropy
- DependenceNonUniformity
- DependenceNonUniformityNormalized
- DependenceVariance
- GrayLevelNonUniformity
- GrayLevelVariance
- HighGrayLevelEmphasis
- LargeDependenceEmphasis
- LargeDependenceHighGrayLevelEmphasis
- LargeDependenceLowGrayLevelEmphasis
- LowGrayLevelEmphasis
- SmallDependenceEmphasis
- SmallDependenceHighGrayLevelEmphasis
- SmallDependenceLowGrayLevelEmphasi

d) Característiques gray level run size zone matrix (GLSZM)

Quantifica les zones de nivell de gris en una imatge, és a dir, el número de vòxels connectats que comparteixen la mateixa intensitat de gris [11]:

- GrayLevelNonUniformity
- GrayLevelNonUniformityNormalized
- GrayLevelVariance
- HighGrayLevelRunEmphasis
- LongRunEmphasis
- LongRunHighGrayLevelEmphasis
- LongRunLowGrayLevelEmphasis
- LowGrayLevelRunEmphasis
- RunEntropy
- RunLengthNonUniformity
- RunLengthNonUniformityNormalized
- RunPercentage
- RunVariance
- ShortRunEmphasis
- ShortRunHighGrayLevelEmphasis
- ShortRunLowGrayLevelEmphasis

e) Característiques gray level run lenght matrix (GLRLM)

Matriu de longitus de'execució del nivell de gris, és a dir, el número de píxels connectats que comparteixen la mateixa intensitat de gris [11]:

- GrayLevelNonUniformity
- GrayLevelNonUniformityNormalized
- GrayLevelVariance
- HighGrayLevelZoneEmphasis
- LargeAreaEmphasis
- LargeAreaHighGrayLevelEmphasis
- LargeAreaLowGrayLevelEmphasis
- LowGrayLevelZoneEmphasis
- SizeZoneNonUniformity
- SizeZoneNonUniformityNormalized
- SmallAreaEmphasis
- SmallAreaHighGrayLevelEmphasis
- SmallAreaLowGrayLevelEmphasis
- ZoneEntropy
- ZonePercentage
- ZoneVariance

f) Característiques neighbouring gray tone difference matrix features (NGTDM)

Una matriu de diferència de l'escala de grisos veïns. Aquestes característiques quantifiquen la diferència entre un valor gris determinat i el valor gris mitjà dels seus veïns [11]:

- Busyness
- Coarseness
- Complexity
- Contrast
- Strength

g) Característiques shape

En aquest grup incloem descriptors de mida i forma 3D y 2D, i que només es calcula sobre la imatge i la màscara no derivada [11]:

- Elongation
- Flatness
- LeastAxisLength
- MajorAxisLength
- Maximum2DDiameterColumn
- Maximum2DDiameterRow
- Maximum2DDiameterSlice
- Maximum3DDiameter
- MeshVolume
- MinorAxisLength
- Sphericity
- SurfaceArea
- SurfaceVolumeRatio
- VoxelVolue

3.4 Característiques clíniques

Com a característiques clíniques d'un pacient entenem els valors referents a símptomes, signes físics, historial mèdic o qualsevols altra informació rellevant relacionada a la condició de salut del nostre pacient. Son característiques avaluades per un especialista mèdic durant el procés de diagnòstic i/o tractament.

Per a l'estudi de radiòmica, aquestes comporten un ajut a l'hora de complementar la informació de la base de dades de la qual farem la predicció, doncs poden afegir dades que, fins i tot, ens ajudi a predir alguna d'aquestes i no només la resposta.

Pel que fa aquest estudi, les característiques clíniques que hem utilitzat són les següents.

Classe	Nombre d'opcions dins cada característica
Histologia	3
Edat	1 (Valor numèric de la edat)
Sexe	2
Tabaquisme	2
Enolisme	2
Resposta	4

Taula 2. Es visualitzen per consola una quantitat de 6 característiques clíniques. Aquestes es poden dividir en diferents classes: Histologia, on diferenciem els adenocarcinomes (ADK: 0), els carcinomes escamosos (CEC: 1), i altres (3); la edat; el sexe, on dividim entre masculí (1) i femení (2); tabaquisme i enolisme, que els diferenciem pel valor 1 (si que fuma o si que veu vi) o 2 (no fuma o no veu vi); la resposta, que pot estar diferenciada en completa (1), parcial (2), estable (3) o progressió (4).

3.5 Models de selecció/reducció de característiques radiòmiques

La selecció de característiques a Python és el procés en que un usuari selecciona de manera automàtica o manual unes característiques del conjunt de dades que més contribueixen amb la variable de predicció o de sortida en la que estem interessats. [12]

Per a reduir la possibilitat de 'overfitting' (Efecte de sobreentrenar un algorisme amb unes certes dades pels quals es coneix el resultat desitjat. En cas de que el nostre model memoritzi i no aprengui, parlarem de 'overfitting'), el número de característiques utilitzades per a la classificació s'ha de reduir a un rang entre 5 i 20. Així llavors, hem utilitzat diferents mètodes per a poder seleccionar característiques radiòmiques i per a reduir la dimensionalitat d'aquestes. [13]

3.5.1 FeatureWiz

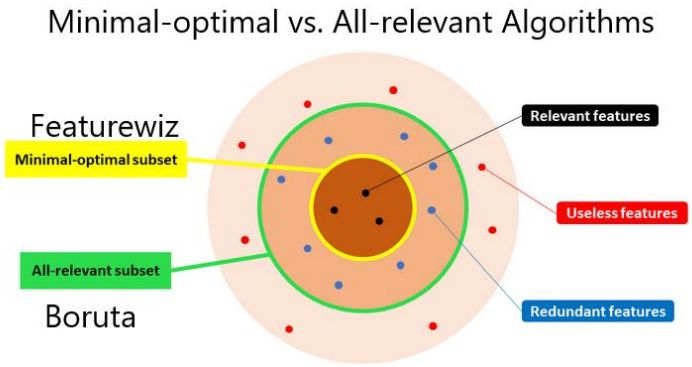


Figura 12. Visualització gràfica del funcionament de la llibreria FeatureWiz.

FeatureWiz és un nou paquet de Python per a crear i seleccionar automàticament característiques importants en un conjunt de dades, fent que es generi el millor model possible per al major rendiment d'aquest. Pot fins i tot veure si el problema és de regressió o de classificació. [14]

Aquest paquet utilitza un algorisme SULOV i Recursive XGBoost per a fer la reducció i selecció de característiques.

Pel que fa a SULOV, aquest fa una cerca de variables no correlacionades realitzant un seguit de passos [15]:

- Troba tots els parells de variables que estan altament correlacionats i que excedeixin un llindar de correlació de 0.8.
- Troba la puntuació de informació mútua per a la variable objectiu. El PIM d'un parell de resultats x e i pertanyent a variables discretes aleatòries X i Y, quantifiquen la diferència entre la probabilitat de la seva coincidència atesa la distribució conjunta i les distribucions individuals, suposant independència matemàtica:

$$\operatorname{pmi}(x;y) \equiv \log \frac{p(x,y)}{p(x)p(y)} = \log \frac{p(x|y)}{p(x)} = \log \frac{p(y|x)}{p(y)}.$$

- Dels parells de variables correlacionades s'eliminen aquelles que tenen un valor de puntuació de informació mútua més baix.
- Es recopilen les variables que tenen el valors de PIM més alts i la menor correlació entre sí.

En canvi, pel que fa a l'algorisme XGBoost recursiu, és un a estratègia en la que s'entrenen models repetidament, eliminant característiques menys importants en cada iteració.

3.5.2 Select K Best

L'anàlisi de variància és una tècnica estadística utilitzada per a analitzar si existeix una diferència significativa entre les mitjanes de tres o més grups de dades podent determinar si aquestes són resultat de l'atzar o realment la diferència és real.

ANOVA es basa en la variància entre els grups i la variància dins de cada grup. Si la variància entre els grups és molt superior a la variància dintre d'aquests, llavors conclourem que hi ha una gran diferència entre les mitjanes dels grups. Per altra banda, si la variància dintre dels grups es molt major a la dels grups per separat, no podrem concloure igual.

La selecció es realitza classificant els valors F i escollint les primeres N característiques, on N és un paràmetre definit per l'usuari

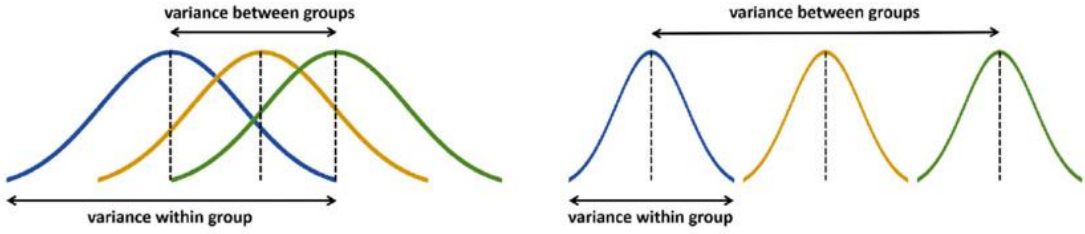


Figura 13. Visualització gràfica de l'anàlisi de variància de Select K Best.

3.6 Models de classificació/predicció

3.6.1 Random Forest (RF)

Parlem d'un algorisme d'aprenentatge automàtic per a tasques supervisades de classificació i regressió. A la fase d'entrenament, el mètode construeix diferents arbres de decisió que realitzen la feina de predicció de manera independent basant-se en subconjunts aleatoris de característiques d'entrada. La predicció final es realitza a través d'una combinació de les prediccions de tots els arbres individuals.

El seu funcionament pas a pas tracta de seleccionar un conjunt que anomenarem mostra d'arrancada. A partir d'aquesta mostra es construeix un arbre. Per a cada node de l'arbre se selecciona una submostra aleatòria dels predictors. D'aquesta manera es selecciona la millor variable predictora i es divideix el node en dos mena de "fills". El procés de construcció de l'arbre continua fins que s'arribi a un criteri d'aturada (en el nostre cas, la profunditat de l'arbre).

Tot el procés anterior es repeteix N vegades, on N es un valor escollit de la quantitat d'arbres que crearem.

Finalment, calculant la predicció mitjana de cada arbre individuals, es prediu la variable objectiu del conjunt sencer.

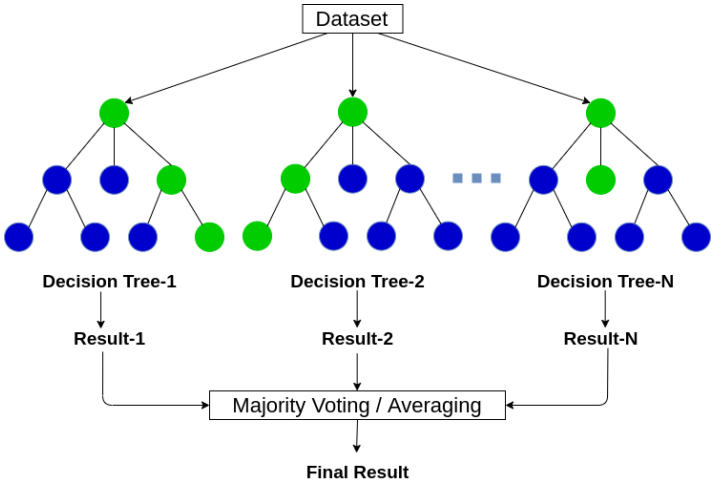


Figura 14. Representació del funcionament d'arbre del classificador Random Forest.

4 Resultats

Disposem d´una col·lecció de 65 pacients amb càncer de pulmó tractats amb SBRT a l'Hospital Universitari Sant Joan de Reus.

De cada pacient tenim la informació corresponent al seu número NHC, que és l'identificador de les històries clíniques hospitalàries [16]; el seu nom complet, amb el nom i dos cognoms; la corresponent imatge TAC simulador amb la estructura 'Smart_pre' verificada per l'especialista mèdic; les característiques radiòmiques extretes de la estructura de cada pacient; la histologia, amb les classes adenocarcinoma (ADK), carcinoma escamós (CEC) i others; la resposta, amb les classes completa, parcial, estable i progressió; i les dades clíniques de cadascun, amb les classes d'edat, sexe, tabaquisme i enolisme.

La nostra variable objectiu ha estat tant la histologia del tumor com la resposta al tractament.

Segons la disposició de les dades que tenim trobem que pel cas de la histologia tenim 33 casos on el tumor és un Adenocarcinoma (ADK), 27 casos on el tumor és un carcinoma escamós (CEC) i 13 casos de la classe others. Com dels 13 últims casos, 8 eren histologies no definides, vam decidir netejar les dades eliminant la classe others, tenint com a objectiu final la predicció de les classes Adenocarcinoma (ADK) i carcinoma escamós (CEC).

Pel que fa a la resposta al tractament 39 casos formen part de la classe de resposta completa, 14 casos per a la resposta parcial, 14 casos per a la resposta estable i 4 casos per a progressió. Com la classe progressió tenia només 4 casos i necessitàvem un numero major per a poder fer un bon entrenament i test, vam decidir netejar el conjunt de dades eliminant aquesta classe. En el cas de la resposta, també es donava que hi havia 3 n/a, així que tant els valors absents com la classe progressió varen quedar eliminats.

Per a una bona pràctica a Python vam assignar a cada classe un valor numèric.

Classe de histologia	**Valor numèric assignat al conjunt de dades**
Adenocarcinoma (ADK)	1
Carcinoma escamós (CEC)	2

Taula 3. Conjunt de classes d'histologia junt amb la codificació a la base de dades.

Classe de resposta	**Valor numèric assignat al conjunt de dades**
Completa	1
Parcial	2
Estable	3

Taula 4. Conjunt de classes de resposta junt amb la codificació a la base de dades.

Per a avaluar els resultats del nostre model, hem fet la prova amb diferents bases de dades, podent fer més endavant la comparativa entre la predicció d'aquestes. Les anomenarem cas 1, cas 2, cas 3, cas 4, cas 5 i cas 6:

- ***Cas 1:*** Característiques radiòmiques (Predir histologia)
- ***Cas 2:*** Característiques radiòmiques (Predir resposta)
- ***Cas 3:*** Variables clíniques i característiques radiòmiques (Predir histologia)
- ***Cas 4:*** Variables clíniques i característiques radiòmiques (Predir resposta)
- ***Cas 5:*** Variables clíniques (Predir histologia)
- ***Cas 6:*** Variables clíniques (Predir resposta)

4.1 Selecció de característiques radiòmiques

Un total de 106 característiques radiòmiques van ser extretes en el procés de segmentació. Després de netejar les dades i utilitzar en primera instància el paquet FeatureWiz, i seguidament la funció Select K Best amb el conjunt de dades de sortida del FeatureWiz, vam simplificar fins a 10 característiques radiòmiques.

Segons el tipus de cas que parlem trobem unes característiques o unes altres.

4.1.1 Característiques finals seleccionades

Casos	Característiques
Cas 1	['LongRunHighGrayLevelEmphasis', '90Percentile', SmallAreaEmphasis', 'SmallDependenceEmphasis', 'Kurtosis', 'Autocorrelation', 'DependenceVariance', 'SurfaceArea', 'SurfaceVolumeRatio', 'HISTOLOGIA'] .
Cas 2	['Elongation', 'DependenceNonUniformityNormalized', 'SurfaceVolumeRatio', 'Maximum3DDiameter', 'Flatness', 'InverseVariance', '90Percentile', 'LargeDependenceEmphasis', 'Maximum2DDiameterSlice', 'RESPUESTA'].
Cas 3	['Idn', 'LongRunHighGrayLevelEmphasis', 'LargeAreaLowGrayLevelEmphasis', 'Imc2', 'SmallAreaEmphasis', 'GrayLevelNonUniformityNormalized.1', 'Maximum', 'SmallAreaHighGrayLevelEmphasis', 'ENOLISMO', 'HISTOLOGIA'].
Cas 4	['DependenceNonUniformityNormalized', 'Elongation', 'Maximum3DDiameter', 'Mean', 'MinorAxisLength', 'DependenceEntropy', 'LargeDependenceEmphasis', '90Percentile', 'Strength', 'RESPUESTA'].
Cas 5	['TABAQUISMO', 'ENOLISMO', 'HISTOLOGIA'].
Cas 6	['SEXO', 'TABAQUISMO', 'ENOLISMO', 'RESPUESTA'].

Taula 5. Característiques radiòmiques seleccionades en cada després d'utilitzar la llibreria FeatureWiz seguit de la funció Select K Best.

4.2 Models de classificació

El model utilitzat és Random Forest junt amb un Cross Validation. En primer moment vèiem que el resultat de la predicció variava una mica segons els cops que el fèiem córrer, així que la validació creuada ens va ajudar a estabilitzar aquest percentatge.

Això succeïa ja que la validació creuada realitza K iteracions on es divideixen les dades d'entrada en K subconjunts. Al fer la repetició i calcular la mitjana aritmètica d'aquestes K iteracions, es pot estimar la precisió del model, en el nostre cas, Random Forest.

Per a veure la millora dels models segons si hem fet el pre-processat de les dades o no, farem la comparativa del Cross Validation amb el model Random Forest amb les 106 característiques inicials i amb les 10 millors característiques seleccionades.

	Predicció sense pre-processat	**Predicció amb pre-processat**	**Desviació estandard de la exactitud**
Cas 1	0.559 (55.9 %)	0.664 (66.4 %)	0.017 (1.7 %)
Cas 2	0.602 (60.2 %)	0.677 (67.7 %)	0.049 (4.9 %)
Cas 3	0.598 (58.9 %)	0.710 (71.0 %)	0.021 (2.1 %)
Cas 4	0.592 (59.2 %)	0.612 (61.2 %)	0.019 (1.9 %)
Cas 5	0.522 (52.2 %)	0.559 (55.9 %)	0.021 (2.1 %)
Cas 6	0.536 (53.6 %)	0.432 (43.2 %)	0.017 (1.7 %)

Taula 6. Comparativa del percentatge d'encert en la predicció entre el model entrenat amb el conjunt de dades sencer o el conjunt de dades després del pre-processament. Càlcul de la desviació estandard de la exactitud de la predicció amb pre-processat.

4.3 Rendiment del nostre estudi

4.3.1 Corba ROC-AUC

La corba ROC-AUC és una eina gràfica amb la que podrem avaluar i comparar el rendiment de models de qualificació com el nostre. Aquesta fa una representació visual de la taxa de vertaders positius en front a la taxa de falsos positius a mesura que es varia el llindar de decisió del model, és a dir, la sensibilitat en front a 1-especificitat.

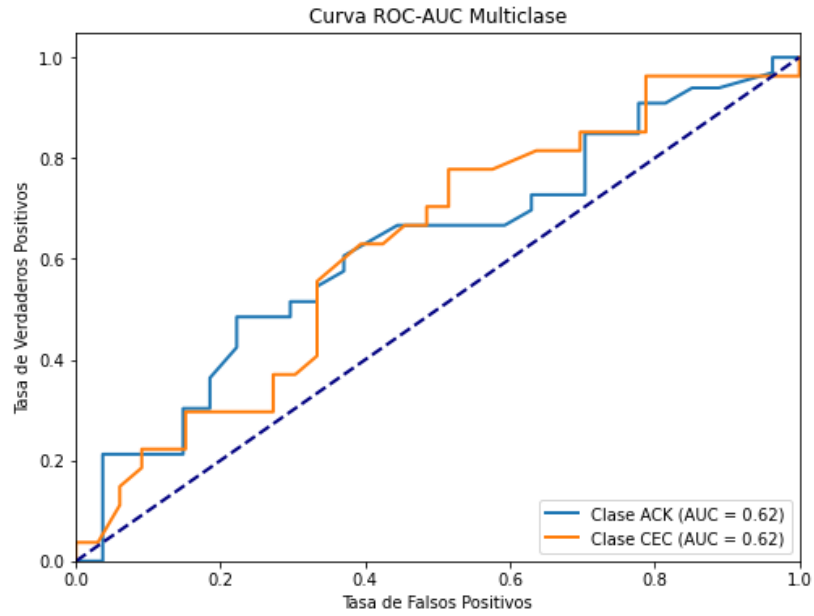
$$\textbf{Sensibilidad (S)} = \frac{\text{Verdaderos Positivos}}{\text{Verdaderos Positivos + Falsos Negativos}} \times 100$$

$$\textbf{Especificidad (E)} = \frac{\text{Verdaderos Negativos}}{\text{Verdaderos Negativos + Falsos Positivos}} \times 100$$

Figura 15. Càlcul de la sensibilitat i especificitat.

Un model ideal d'aquesta corba trobaria una gràfica que s'aproparia la cantonada superior esquerra on indicaria que tant la sensibilitat com la especificitat és 1, referint-se a que no hi ha falsos positius i tots els vertaders positius s'han capturat.

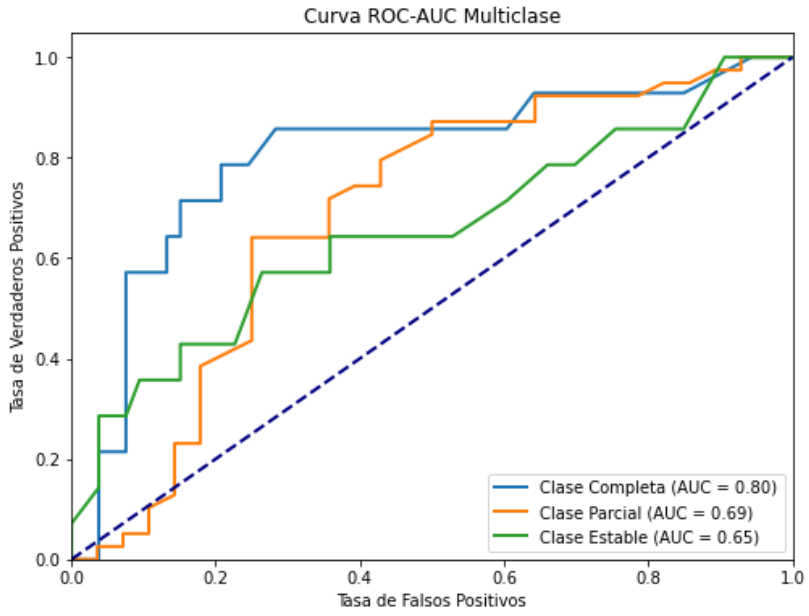
4.3.1.1 Cas 1




```
AUC Macro: 0.62
```

Figura 16. Corba AUC-ROC per a la histologia del cas 1. Adenocarcinomes (blau), carcinomes escamosos (taronja).

4.3.1.2 Cas 2




```
AUC Macro: 0.71
```

Figura 17. Corba AUC-ROC per a la resposta del cas 2. Completa (blau), parcial (taronja), estable (verd)

4.3.1.3 Cas 3

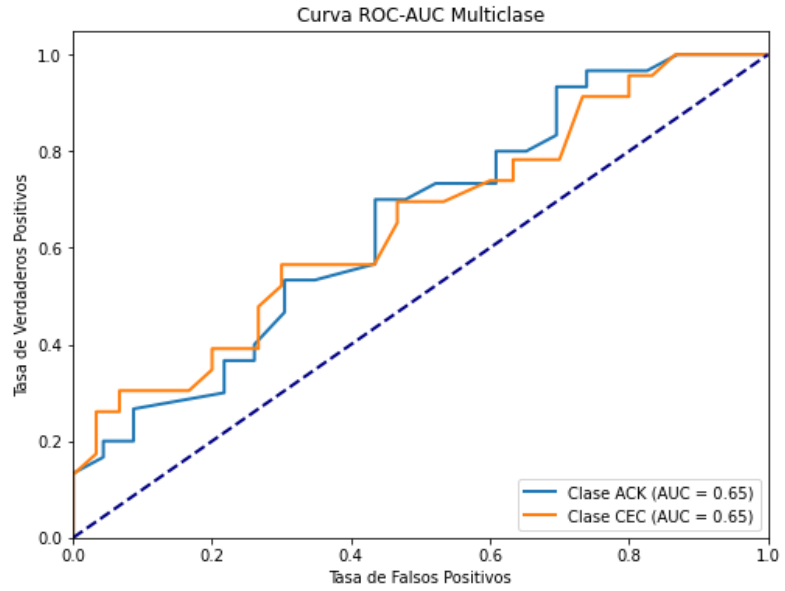


Figura 18. Corba AUC-ROC per a la histologia del cas 1. Adenocarcinomes (blau), carcinomes escamosos (taronja).

4.3.1.4 Cas 4

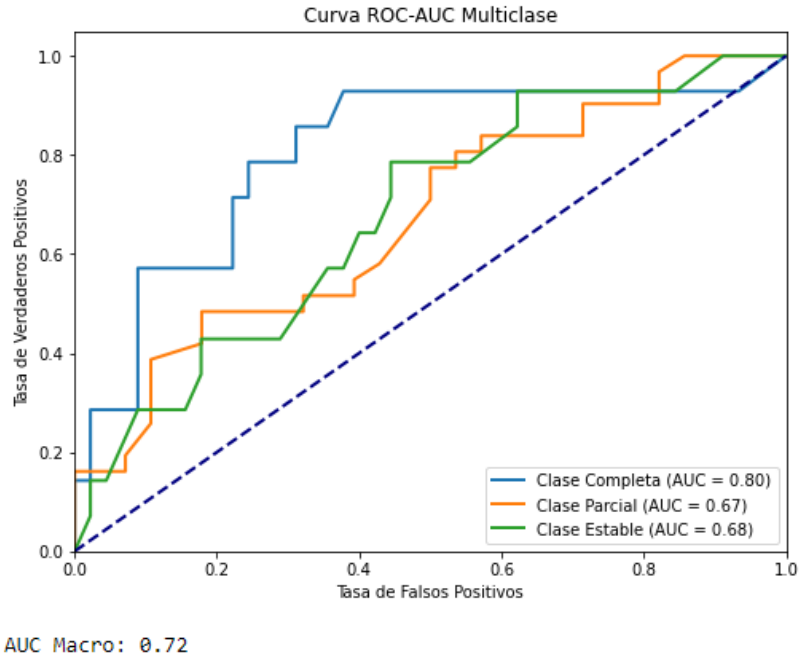
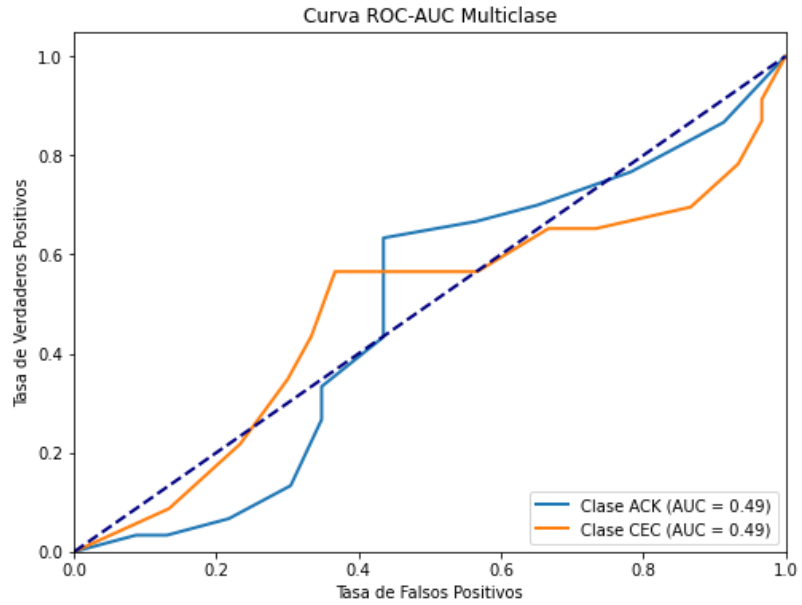


Figura 19. Corba AUC-ROC per a la resposta del cas 2. Completa (blau), parcial (taronja), estable (verd)

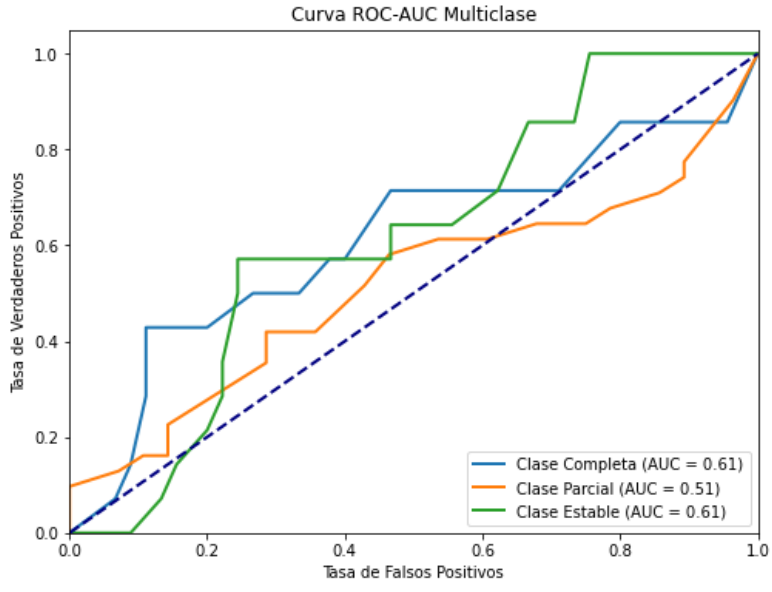
4.3.1.5 Cas 5




```
AUC Macro: 0.49
```

Figura 20. Corba AUC-ROC per a la histologia del cas 1. Adenocarcinomes (blau), carcinomes escamosos (taronja).

4.3.1.6 Cas 6




```
AUC Macro: 0.58
```

Figura 21. Corba AUC-ROC per a la resposta del cas 2. Completa (blau), parcial (taronja), estable (verd)

4.3.2 Taula de confusió

La taula de confusió és una matriu que ens permet veure els encerts i els errors del nostre model envers a una classificació/predicció. Per a entendre el seu funcionament hem de comprendre que a la matriu es presentaran tots els casos, tants els encertats com els erronis de la predicció [8]. La disposició a la matrius serà del següent tipus:

Figura 22. Característiques generals de la matriu de confusió.

En els nostres casos, ja sigui per a la predicció de la histologia o de la resposta, veurem que els valors en diagonal que aquí es disposen de color verd, seran les prediccions correctes de les diferents classes d'histologia (ADK, CEC, others), o l'encert en la predicció de la resposta al tractament (Completa, Parcial, Estable, Progressió).

4.3.2.1 Cas 1

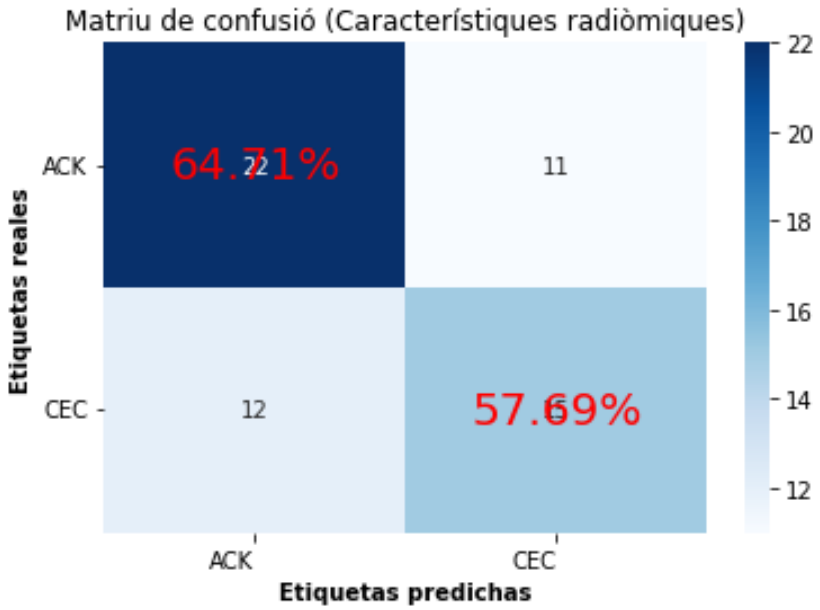


Figura 23. Percentatge d'encert amb variables radiòmiques de les diferents classes d'histologia (ADK : 1, CEC: 2).

4.3.2.2 Cas 2

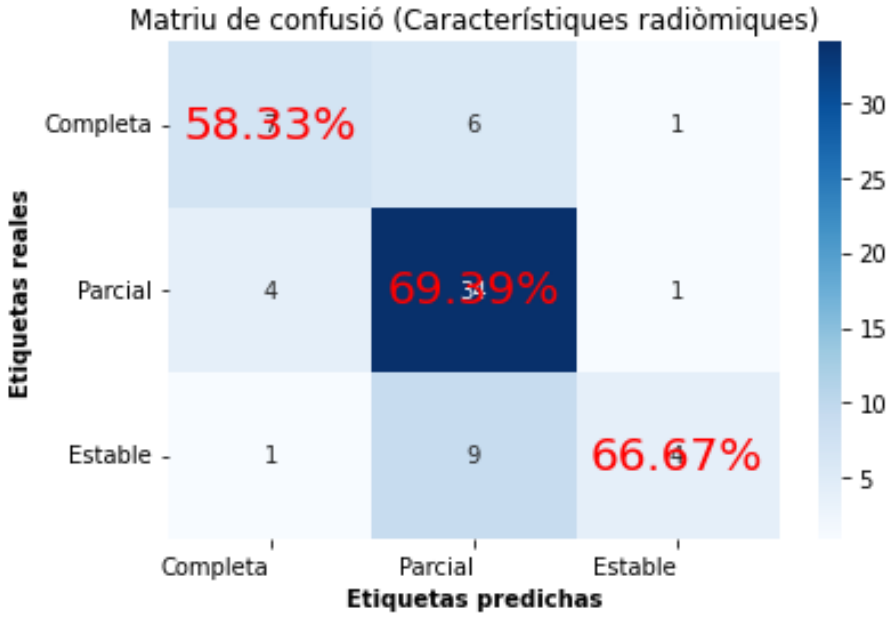


Figura 24. Percentatge d'encert de la resposta al tractament amb variables radiòmiques (la resposta, que pot estar diferenciada en completa (1), parcial (2), estable (3)).

4.3.2.3 Cas 3

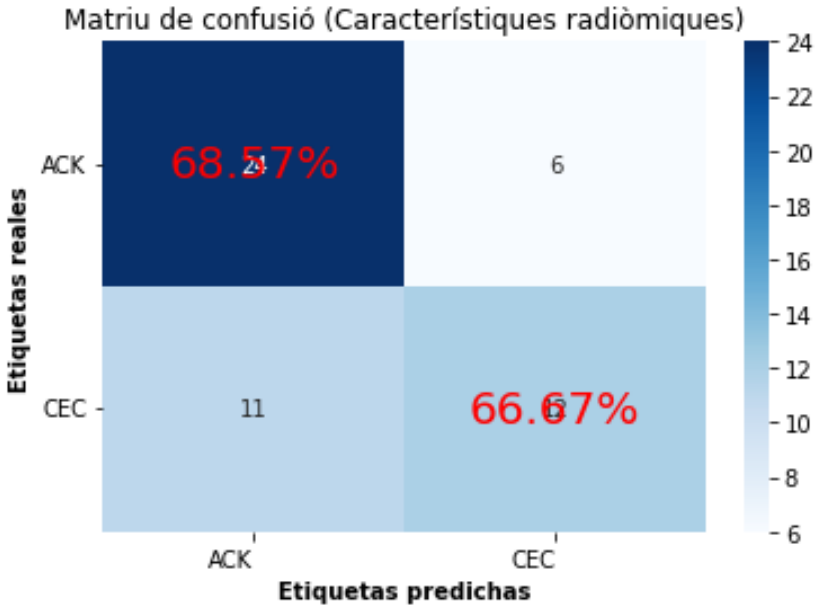


Figura 25. Percentatge d'encert amb variables radiòmiques i clíniques dels diferents classes d'histologia (ADK: 1, CEC: 2).

4.3.2.4 Cas 4

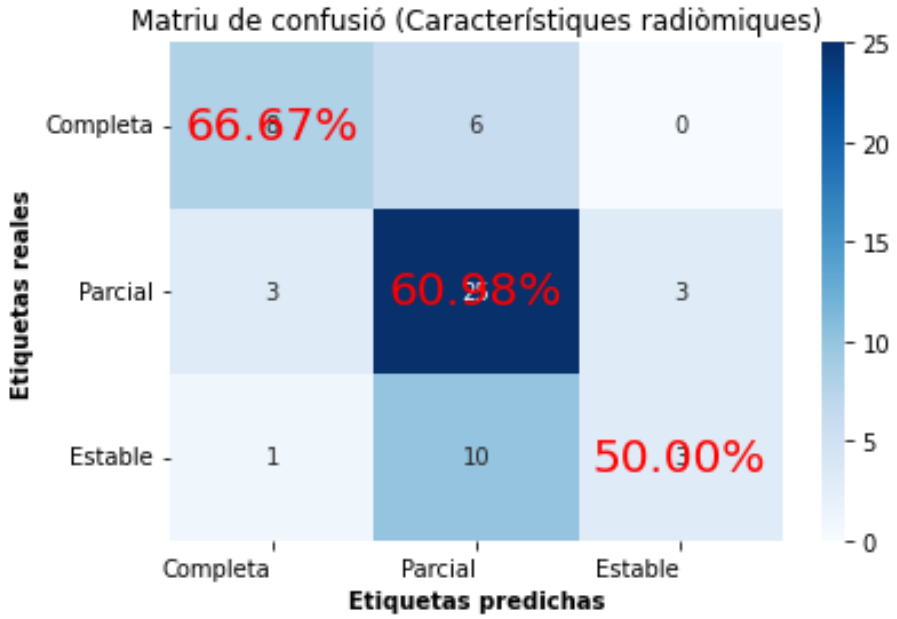


Figura 26. Percentatge d'encert de la resposta al tractament amb variables radiòmiques i clíniques (la resposta, que pot estar diferenciada en completa (1), parcial (2), estable (3)).

4.3.2.5 Cas 5

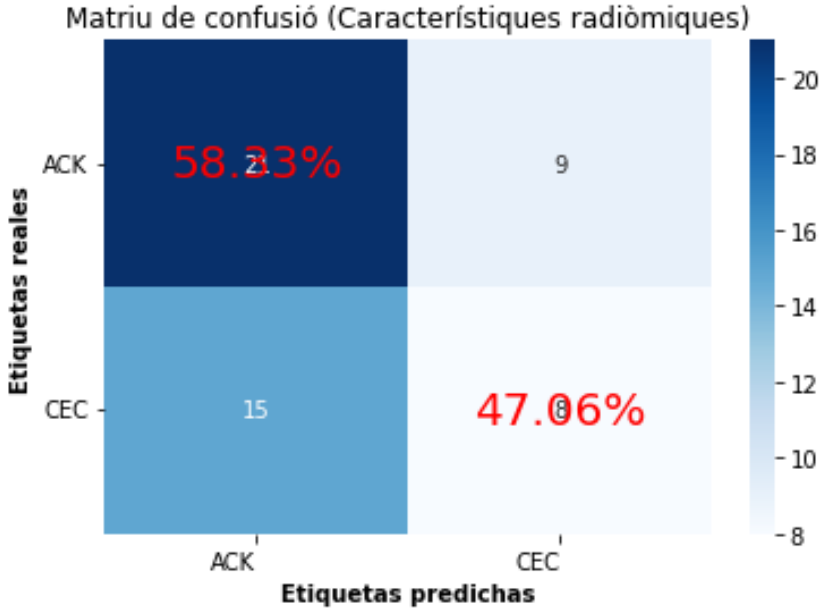


Figura 27. Percentatge d'encert amb variables clíniques dels diferents classe d'histologia (ADK: 1, CEC: 2).

4.3.2.6 Cas 6

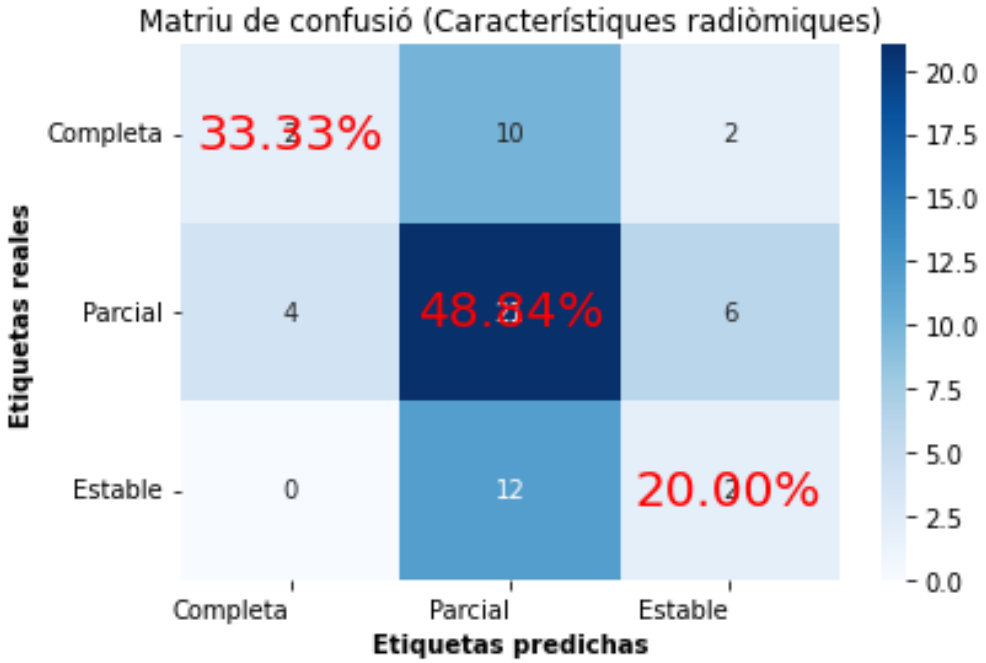


Figura 28. Percentatge d'encert de la resposta al tractament amb variables clíniques (la resposta, que pot estar diferenciada en completa (1), parcial (2), estable (3)).

5 Discussió

L'anàlisi de característiques radiòmiques derivades de imatges del TAC pre ha estat utilitzat per a elaborar un model predictiu de la resposta i de la histologia per a obtenir un nivell de precisió superior a qualsevol predicció amb un anàlisi radiològic convencional.

El model predictiu era una classificació amb Random Forest. La selecció de característiques que usaríem al model eren la sortida de la llibreria FeatureWiz junt amb la funció Select K Best (ANOVA), on la llibreria FeatureWiz implementava el mètode SULOV i XGBoost

En aquest treball també, hem implementat i investigat diferents enfocaments havent creat 3 conjunts de dades dels pacients observats en el cas de resposta i 3 per a la histologia. Pels dos primers casos vam generar un conjunt basat en característiques radiòmiques extretes de Slicer 3D de cadascun dels TAC pre dels pacients; pel segons, un conjunt relacionant característiques radiòmiques junt amb les dades clíniques; i finalment, 2 conjunts on només fèiem ús del conjunt de dades clíniques.

La síntesis del treball anava guiada per a diferents passos on extrauríem les característiques radiòmiques, crearíem el conjunt de dades, utilitzaríem una combinació de mètodes de selecció amb un mètode de classificació considerant una reducció de fins a 10 característiques finals que usaríem com a possibles marcadors de l'objectiu proposat.

Encara que en els diferents casos els resultats acostumen a millorar un cop processades les dades, considero que hi ha casos amb rendiment molt baix. Si fem una primera observació a les corbes AUC-ROC podem veure que hi ha casos en que no es fa de manera correcta una

separació dels valors de vertaders negatius, falsos negatius, falsos positius i vertaders positius. Pel cas 1, 3, 5 i 6 veiem que la AUC-ROC CURVE és inferior a 0.7. Això ens indica que el model no distingeix bé entre les classes i que pot predir la classe positiva com a negativa i viceversa.

Si fem una ullada a les matrius de confusió podem entendre millor això, i es que pel cas de la predicció de la histologia, els percentatges d'encert s'apropen a al percentatge de les AUC-ROC Curve. Aquests resultats van al voltant de entre un 60% i un 70% pels casos 1 i 3, mentre que el cas 5 el resultats no són gens significatius. Era d'esperar que els conjunt de dades clíniques no tingués una bona predicció, doncs les variables que tenim no només són un nombre molt petit, sinó que són massa genèriques i no impliquen cap tipus d'importància. El que si que s'observa és que la predicció dels adenocarcinomes (ADK) acostuma a ser millor que la dels carcinomes escamosos (CEC). Això es deu donar a causa de que el nombre de casos de la classe ADK és major al de CEC, proporcionant així un millor entrenament, ja que en altres estudis [18] s'ha pogut veure que els carcinomes escamosos tenen millor relació amb la radiòmica.

Enfocant-nos a la resposta veiem que té una dinàmica similar a la histologia quan només seleccionem les dades clíniques, i es que no tenen cap tipus de relació amb la predicció, fent que no hi hagi una bona discriminació entre classes i cap tipus de millora. Podem considerar que fa prediccions a l'atzar. Això canvia en els casos 2 i 4. Les AUC-ROC curves tenen una mitja superior al 0.7 indicant que el model pot distingir entre classes, i les prediccions d'alguna de les classes és propera al 70%. La diferència entre aquests 2 casos és que pel cas 2, les classes amb millor predicció no són les mateixes que en el cas 4. Amb això podem observar que la selecció de característiques radiòmiques del cas 4 sigui més encertada o més significativa que en el cas 2, ja que les dades clíniques han demostrat, segons el funcionament de la llibreria FeatureWiz, que hi ha variables amb poca variància o importància que no s'han eliminat o detectat en el cas de la radiòmica. La classe 'Completa' té un nombre de casos menor que la classe 'Parcial' fent que a priori, i com si succeeix al cas 2, tingui una pitjor predicció que l'anterior. Però no és així. A part, la AUC-ROC curve de la classe 'Completa' és superior al 0.8 en ambdós casos, donant a entendre que és la classe que millor discriminació té i que prediu millor amb les característiques seleccionades del cas 4.

En general podem dir que en el cas de la predicció d'histologia hi ha un marge de millora molt gran. Doncs realment no he estat capaç de que hi hagi una discriminació ideal entre classes per a que el model sigui capaç de entrenar-se correctament i distingir-les. Pel referent a la resposta, hem vist els resultats positius del cas 2 i 4, donant més importància a la classe 'Completa' que és la que millor AUC-ROC curve presenta.

També podem veure que les nostres variables clíniques han ajudat en els casos que estan en el mateix conjunt que les característiques radiòmiques, doncs la predicció i la discriminació entre classes era millor que per separat encara que les característiques radiòmiques són les que tenen més pes.

Això mateix ho he comparat amb altres investigacions similars. A l'article 'Radiomics and deep learning methods for the prediction of 2-year overall survival in LUNG1 dataset' [13], veiem com després de tenir un conjunt de dades clíniques i una altra amb radiòmiques , de fer una selecció amb 4 mètodes diferents i una comparativa amb 6 models, arriben a la conclusió de que les característiques radiòmiques són més significatives que les dades clíniques. A part, ho fan utilitzant el model de Random Forest amb la selecció amb ANOVA que, segons els mateix article, és la millor combinació per a aquests casos.

A l'article 'Selección de características con método wrapper para un sistema de detección de intruso: caso CICIDS-2017' també mostra com per a problemes de classificació, el model de Random Forest que hem utilitzat és el més precís. [17]

Altra part important a tenir en compte és la aposta per l'ús de la llibreria FeatureWiz i la funció Select K Best com a mètode de selecció de característiques, doncs encara que realitza el que volíem i necessitàvem per a processar les dades, fins a obtenir els resultats no hem pogut valorar el rendiment. FeatureWiz és un mètode prou ràpid i representatiu que indica en tot el procediment les passes que realitza i els valors que té com a sortida durant el procés. Select K Best implementa mètodes coneguts com el ANOVA selector, que com hem vist en alguns articles [16], és el més adient per a aquest tipus d'investigació. Encara i així, hem vist com hem tingut problema de discriminació d'algunes classes, sobretot en la predicció d'histologia. Pels diferents casos, sempre hi sortien com a possibles marcadors unes 10 característiques radiòmiques amb bastantes variacions, fent que realment no hi hagi un patró o cap tipus de importància significativa d'aquestes. Encara i així, s'observa una millora de la predicció i discriminació amb les seleccions de característiques dels conjunts de dades amb característiques radiòmiques i dades clíniques alhora.

Fent la comparativa amb el TFG anterior a aquest veiem que només hi ha 1 característica radiòmica que coincideix com a important. Aquesta és '90Percentile', que junt amb 'Elongation', 'DependenceNonUniformityNormalized', 'Maximum3DDiameter' i 'LargeDependenceEmhasis' (Coinicideixen com a top 10 variables tant en el cas 2 com 4) han aconseguit discriminar la classe 'Completa' i predir-la amb una mica més d'encert que les altres.

6 Conclusió

Un cop finalitzat el treball podem dir que s'han complert els objectius plantejats ja que tenim un conjunt de dades més ampli que permet que el model s'entreni i no memoritzi les dades, evitant resultats ficticis que realment no demostren que serveixin com ajuda al diagnòstic. També hem refinat l'extracció de dades radiòmiques gràcies a la smart segmentation. Al treball anterior aquesta segmentació es feia amb la pròpia plataforma de 3D Slicer de manera manual, a ull. Ara és un algorisme refinat que ens ajuda a poder seleccionar un volum més ajustat de la lesió, aconseguint que les delimitació més ajustada proporcioni uns valors més exactes de les característiques radiòmiques. Pel que fa als models predictius, hem donat total importància a Random Forest aconseguint que en alguns casos particulars, es vegi potencial de predicció. Ara també, les variables objectiu eren diferents a les del treball anterior, doncs no només ens centrem en les metàstasis, sinó que valorem la possibilitat d'ampliar la predicció a tots els casos possibles d'histologia i de resposta que tenim. Això fa que el percentatge de predicció no sigui molt significatiu però si que ens pugui servir com a referència de per on ha de seguir i enfocar-se el treball.

En referència a possibles biomarcadors concloem en que encara no hi ha una referència clara de quines característiques radiòmiques de 3D Slicer, doncs només hi ha 1 característica que coincideixi amb el treball anterior.

Finalment podem dir que la aplicació clínica d'aquest resultats no és viable en aquests moments, encara que amb una mica més de desenvolupament podrà aconseguir uns resultats suficients com per a funcionar com a eina d'ajut a l'especialista mèdic.

Com a seguit d'aquest treball crec que s'ha de donar més emfasi a la barreja entre dades clíniques i característiques radiòmiques, doncs en els nostres casos, ens ha servit per a discriminar millor les classes. L'ampliació de dades clíniques per a exploracions de pacients amb lesions pulmonars tractats amb SBRT ha de ser el pas a seguir.

Un cop això seguir provant mètodes de selecció i/o reducció diferents podrà servir com a comparativa per a valorar la importància real de diferents característiques radiòmiques.

7 Referències

[1] "El cáncer de pulmón aumenta su letalidad en España - GECP". GECP. Accedido el 5 de septiembre de 2023. [En línea]. Disponible: https://www.gecp.org/el-cancer-de-pulmon-aumenta-su-letalidad-en-espana/

[2] "Radiocirugía estereotáctica (SRS/SBRT)". Radiologyinfo.org. Accedido el 21 de mayo de 2023. [En línea]. Disponible: https://www.radiologyinfo.org/es/info/stereotactic

[3] "TrueBeam | varian". Home | Varian. Accedido el 5 de septiembre de 2023. [En línea]. Disponible: https://www.varian.com/es/products/radiotherapy/treatment-delivery/truebeam

[4] R. J. Gillies, P. E. Kinahan y H. Hricak, "Radiomics: Images are more than pictures, they are data", *Radiology*, vol. 278, n.º 2, pp. 563–577, febrero de 2016. doi: 10.1148/radiol.2015151169.

[5] O. Diaz, *0.2 X-Ray Imaging (Part 2)*, 2ª ed. Master Biomedical Data Science.

[6] Y. Liu *et al.*, "Imaging biomarkers to predict and evaluate the effectiveness of immunotherapy in advanced non-small-cell lung cancer", *Frontiers Oncol.*, vol. 11, marzo de 2021. doi: 10.3389/fonc.2021.657615.

[7] M. Avanzo, J. Stancanello, G. Pirrone y G. Sartor, "Radiomics and deep learning in lung cancer", *Strahlentherapie und Onkologie*, vol. 196, n.º 10, pp. 879–887, mayo de 2020. doi: 10.1007/s00066-020-01625-9.

[8] "SlicerRT". SlicerRT. Accedido el 6 de septiembre de 2023. [En línea]. Disponible: https://slicerrt.github.io/

[9] "Sample data — 3D slicer documentation". Welcome to 3D Slicer's documentation! — 3D Slicer documentation. Accedido el 5 de septiembre de 2023. [En línea]. Disponible: https://slicer.readthedocs.io/en/5.2/user_guide/modules/sampledata.html

[10] R. J. Gillies, P. E. Kinahan y H. Hricak, "Radiomics: Images are more than pictures, they are data", *Radiology*, vol. 278, n.º 2, pp. 563–577, febrero de 2016. doi: 10.1148/radiol.2015151169.

[11] "Radiomics". Radiomics. Accedido el 5 de septiembre de 2023. [En línea]. Disponible: https://www.radiomics.io/slicerradiomics.html

[12] Manuel Terradez Gurrea, "Anaálisis de componentes principales", resumen extendido de, UOC.

[13] D. David. "Selección automática de características en Python: Una guía esencial | HackerNoon". HackerNoon - read, write and learn about any technology. Accedido el 5 de septiembre de 2023. [En línea]. Disponible: https://hackernoon.com/es/seleccion-automatica-de-caracteristicas-en-python-una-guia-esencial-uv3e37mk

[14] "GitHub - autoviml/featurewiz: Use advanced feature engineering strategies and select best features from your data set with a single line of code." GitHub. Accedido el 6 de septiembre de 2023. [En línea]. Disponible: https://github.com/AutoViML/featurewiz

[15] D. David. "Selección automática de características en Python: Una guía esencial | HackerNoon". HackerNoon - read, write and learn about any technology. Accedido el 6 de septiembre de 2023. [En línea]. Disponible: https://hackernoon.com/es/seleccion-automatica-de-caracteristicas-en-python-una-guia-esencial-uv3e37mk

[16] Unknown. "¿Qué es el NHC o número de historia clínica?" HAY UNA GRAN DIFERENCIA ENTRE HAZ LO POSIBLE Y HAZLO POSIBLE. Accedido el 5 de septiembre de 2023. [En línea]. Disponible: http://asesfive.blogspot.com/2018/05/que-es-el-nhc-o-numero-de-historia.html

[17] A. Braghetto, F. Marturano, M. Paiusco, M. Baiesi y A. Bettinelli, "Radiomics and deep learning methods for the prediction of 2-year overall survival in LUNG1 dataset", *Scientific Rep.*, vol. 12, n.º 1, agosto de 2022. doi: 10.1038/s41598-022-18085-z.

8 Annexos

8.1 Annex 1: Codi del programa

8.1.1 Importació de les dades

```
# import packages
import seaborn as sns
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import seaborn as sns
from sklearn import svm
import xgboost as xgb
import sklearn.metrics
from sklearn.model_selection import StratifiedKFold, KFold
from sklearn.metrics import roc_curve, auc
from sklearn.metrics import confusion_matrix
from sklearn.model_selection import train_test_split, cross_val_score, cross_val_predict
from sklearn.preprocessing import StandardScaler
from sklearn.ensemble import RandomForestClassifier, GradientBoostingClassifier
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import accuracy_score, roc_auc_score, confusion_matrix
from featurewiz import featurewiz
from sklearn.feature_selection import SelectKBest, f_classif
```

Importació de les dades

```
#Importar dades
data = pd.read_excel('PRUEBA_RESPUESTA_clinicas_radiomica.xlsx')
data = data.set_index('NumTAC')
# Eliminar valores ausentes de la Columna1
data = data.dropna()
data = data.drop(data[data['RESPUESTA'] == 4].index)
X = data.drop('RESPUESTA', axis=1)
y = data.RESPUESTA.values

#Estandarice las características usando StandardScaler de scikit-learn.
X_scaled = StandardScaler().fit_transform(X)

# Convertir el arreglo en una serie de pandas
serie = pd.Series(y)
# Utilizar el método value_counts() en la serie
conteo = serie.value_counts()
print(conteo)
data.head()
```

```
2.0    31
3.0    14
1.0    14
Name: count, dtype: int64
```

NumTAC	RESPUESTA	Elongation	Flatness	LeastAxisLength	MajorAxisLength	Maximum2DDiameterColumn	Maximum2DDiameterRow	Maximum2DDiameterSlic
1	2.0	0.763887	0.612364	12.224396	19.962643	20.899723	23.257148	17.26748
2	3.0	0.789238	0.713906	13.900057	19.470426	22.388988	20.899723	21.73087
3	2.0	0.785386	0.632430	9.037082	14.289449	14.638016	14.253048	16.05843
5	2.0	0.701980	0.591749	17.725316	29.954129	31.426653	31.437813	25.54881
6	3.0	0.795386	0.589448	15.320899	25.991963	26.475150	33.328121	25.39062

5 rows × 112 columns

8.1.2 Primera predicció sense pre-processat

Sense validació creuada amb Random Forest

```
#Dividim entre entrenaent i validacio
X_train, X_valid, y_train, y_valid = train_test_split(X_scaled,y,test_size = 0.2 ,stratify=y, random_state = 0, shuffle = True)
classifier = RandomForestClassifier(n_estimators = 100, criterion = 'gini', max_depth = None,
                                    min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0,
                                    max_features = 'sqrt', max_leaf_nodes = None, min_impurity_decrease = 0.0,
                                    bootstrap = True, oob_score = True, n_jobs = None, random_state = None, verbose = 0,
                                    warm_start = False, class_weight = None, ccp_alpha = 0.0)

classifier.fit(X_train,y_train)
# make prediction
preds = classifier.predict(X_valid)
# check performance
accuracy_score(y_valid, preds)
```

```
0.5833333333333334
```

Amb validació creuada. Splits = 5

```
# Inicializar el validador cruzado (Stratified K-Fold)
n_splits = 5
skf = StratifiedKFold(n_splits=n_splits, shuffle=True, random_state=42)

# Inicializar una lista para almacenar los resultados
accuracy_scores = []
all_y_true = []
all_y_pred = []

# Ciclo de validación cruzada
for train_index, test_index in skf.split(X_scaled, y):
    X_train, X_val = X_scaled[train_index], X_scaled[test_index]
    y_train, y_val = y[train_index], y[test_index]

    # Crear el modelo de Random Forest Classifier
    rf_classifier = RandomForestClassifier(n_estimators = 100, criterion = 'gini', max_depth = None,
                                    min_samples_split = 2, min_samples_leaf = 1, min_weight_fraction_leaf = 0.0,
                                    max_features = 'sqrt', max_leaf_nodes = None, min_impurity_decrease = 0.0,
                                    bootstrap = True, oob_score = True, n_jobs = None, random_state = None, verbose = 0,
                                    warm_start = False, class_weight = None, ccp_alpha = 0.0)
    # Entrenar el modelo en el conjunto de entrenamiento
    rf_classifier.fit(X_train, y_train)

    # Realizar predicciones en el conjunto de prueba
    y_pred = rf_classifier.predict(X_val)
      # Almacenar las etiquetas reales y las predicciones
    all_y_true.extend(y_val)
    all_y_pred.extend(y_pred)

    # Calcular la exactitud y almacenarla en la lista
    accuracy = accuracy_score(all_y_true, all_y_pred)
    accuracy_scores.append(accuracy)

# Calcular la exactitud promedio y desviación estándar de las puntuaciones
average_accuracy = np.mean(accuracy_scores)
std_accuracy = np.std(accuracy_scores)

print("Exactitud promedio:", average_accuracy)
print("Desviación estándar de la exactitud:", std_accuracy)
print(y_val)
```

```
Exactitud promedio: 0.6090866290018833
Desviación estándar de la exactitud: 0.036640256018284016
[2. 3. 1. 2. 1. 2. 2. 2. 3. 3. 2.]
```

8.1.3 Selecció/Reducció de característiques radiòmiques

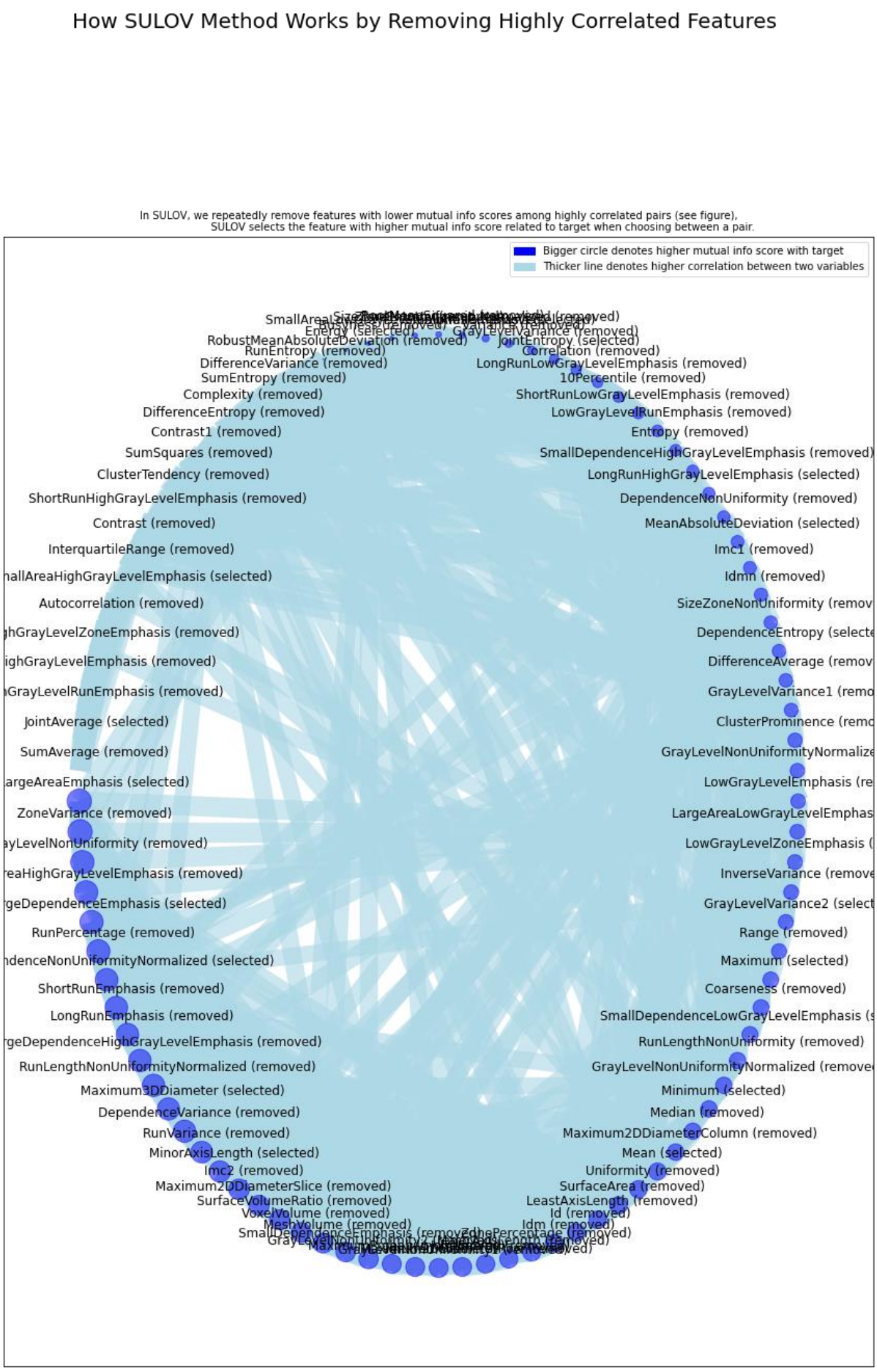
8.1.3.1 LLibreria FeatureWiz

Reducció de característiques radiòmiques

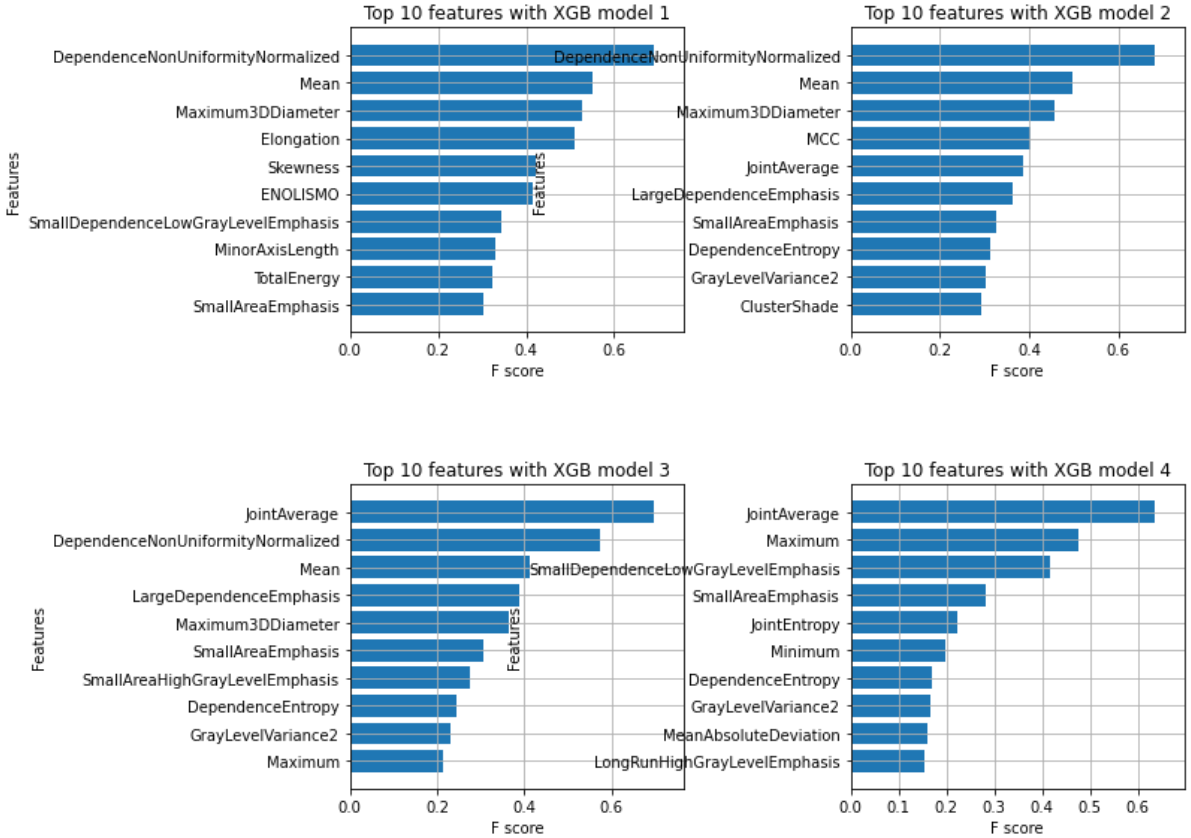
Llibreria featurewiz

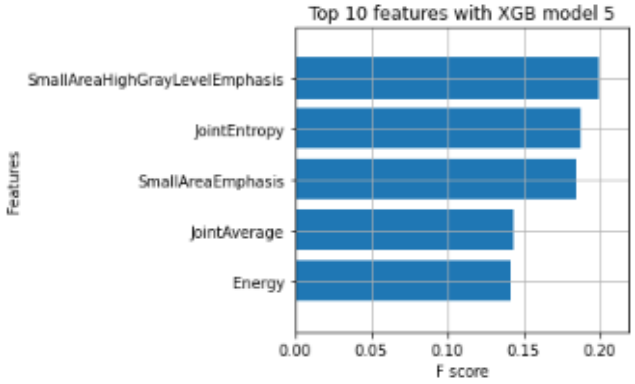
```
#Reduccio de variables
# automatic feature selection by using featurewiz package
features, train = featurewiz(data, target='RESPUESTA', corr_limit= 0.8 , verbose= 2 , sep= "," , header= 0 ,
                             test_data= "" , feature_engg= "" , category_encoders= "", dask_xgboost_flag=False,
                             nrows=None, skip_sulov=False, skip_xgboost=False)

#Reduccio de variables2
# automatic feature selection by using featurewiz package
features, train2 = featurewiz(train, target='RESPUESTA', corr_limit= 0.8 , verbose= 2 , sep= "," , header= 0 ,
                              test_data= "" , feature_engg= "" , category_encoders= "", dask_xgboost_flag=False,
                              nrows=None, skip_sulov=False, skip_xgboost=False)
```

```
############################################################################################
############      F A S T   F E A T U R E  E N G G   A N D   S E L E C T I O N ! ########
# Be judicious with featurewiz. Don't use it to create too many un-interpretable features! #
############################################################################################
featurewiz has selected 0.8 as the correlation limit. Change this limit to fit your needs...
Skipping feature engineering since no feature_engg input...
Skipping category encoding since no category encoders specified in input...
#### Single_Label Multi_Classification problem ####
    Loaded train data. Shape = (59, 112)
    Some column names had special characters which were removed...
#### Single_Label Multi_Classification problem ####
No test data filename given...
############################################################################################
######################## C L A S S I F Y I N G  V A R I A B L E S  ###################
############################################################################################
        No variables were removed since no ID or low-information variables found in data set
No GPU active on this device
    Tuning XGBoost using CPU hyper-parameters. This will take time...
Removing 0 columns from further processing since ID or low information variables
    After removing redundant variables from further processing, features left = 111
No interactions created for categorical vars since feature engg does not specify it
    target labels need to be converted...
Completed label encoding of target variable = RESPUESTA
How model predictions need to be transformed for RESPUESTA:
        {0: 1.0, 1: 2.0, 2: 3.0}
############################################################################################
#####  Searching for Uncorrelated List Of Variables (SULOV) in 111 features ############
############################################################################################
    there are no null values in dataset...
    Removing (78) highly correlated variables:
```

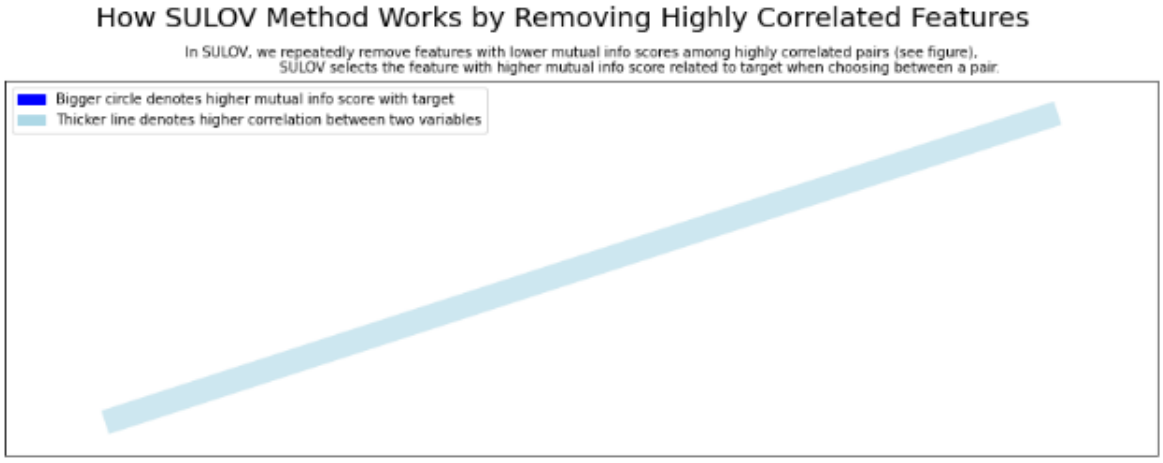

How SULOV Method Works by Removing Highly Correlated Features
In SULOV, we repeatedly remove features with lower mutual info scores among highly correlated pairs (see figure),
SULOV selects the feature with higher mutual info score related to target when choosing between a pair.
Bigger circle denotes higher mutual info score with target
Thicker line denotes higher correlation between two variables
RobustMeanAbsoluteDeviation (removed)
RunEntropy (removed)
DifferenceVariance (removed)
SumEntropy (removed)
Complexity (removed)
DifferenceEntropy (removed)
Contrast1 (removed)
SumSquares (removed)
ClusterTendency (removed)
ShortRunHighGrayLevelEmphasis (removed)
Contrast (removed)
InterquartileRange (removed)
Autocorrelation (removed)
JointAverage (selected)
SumAverage (removed)
ZoneVariance (removed)
RunPercentage (removed)
ShortRunEmphasis (removed)
LongRunEmphasis (removed)
RunLengthNonUniformityNormalized (removed)
Maximum3DDiameter (selected)
DependenceVariance (removed)
RunVariance (removed)
MinorAxisLength (selected)
Imc2 (removed)
Maximum2DDiameterSlice (removed)
SurfaceVolumeRatio (removed)
VoxelVolume (removed)
MeshVolume (removed)
JointEntropy (selected)
Correlation (removed)
LongRunLowGrayLevelEmphasis (removed)
10Percentile (removed)
ShortRunLowGrayLevelEmphasis (removed)
LowGrayLevelRunEmphasis (removed)
Entropy (removed)
LongRunHighGrayLevelEmphasis (selected)
DependenceNonUniformity (removed)
MeanAbsoluteDeviation (selected)
Imc1 (removed)
Idmn (removed)
DependenceEntropy (selecte
DifferenceAverage (remov
GrayLevelVariance1 (remo
ClusterProminence (remo
LowGrayLevelEmphasis (re
LowGrayLevelZoneEmphasis (
InverseVariance (remove
GrayLevelVariance2 (select
Range (removed)
Maximum (selected)
Coarseness (removed)
RunLengthNonUniformity (removed)
Minimum (selected)
Median (removed)
Maximum2DDiameterColumn (removed)
Mean (selected)
Uniformity (removed)
SurfaceArea (removed)
LeastAxisLength (removed)
Id (removed)
Idm (removed)

```
Time taken for SULOV method = 2 seconds
    Adding 0 categorical variables to reduced numeric variables  of 33
Finally 33 vars selected after SULOV
Converting all features to numeric before sending to XGBoost...
###############################################################################
#####    R E C U R S I V E   X G B O O S T : F E A T U R E   S E L E C T I O N  #######
###############################################################################
    using regular XGBoost
    Taking top 22 features per iteration...
    XGBoost version using 1.7.5 as tree method: hist
    Number of booster rounds = 100
        Selected: ['DependenceNonUniformityNormalized', 'Elongation', 'SmallAreaEmphasis', 'Maximum3D
Diameter', 'Mean', 'MinorAxisLength', 'MeanAbsoluteDeviation', 'Flatness', 'Sphericity', 'GrayLevelVa
riance2', 'Maximum', 'Skewness', 'DependenceEntropy', 'LargeDependenceEmphasis', 'LargeDependenceLowG
rayLevelEmphasis', 'TotalEnergy', 'Minimum', 'Kurtosis', 'EDAD', 'MCC', '90Percentile', 'ClusterShad
e']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['DependenceNonUniformityNormalized', 'Maximum3DDiameter', 'SmallAreaEmphasis', 'Me
an', 'LargeDependenceEmphasis', 'DependenceEntropy', 'Maximum', 'ClusterShade', 'GrayLevelVariance2',
'Minimum', 'MCC', 'MeanAbsoluteDeviation', 'LargeDependenceLowGrayLevelEmphasis', 'LongRunHighGrayLev
elEmphasis', 'MinorAxisLength', 'Energy', 'SEXO', 'JointAverage', 'TABAQUISMO', 'EDAD', 'SmallAreaHig
hGrayLevelEmphasis', 'Strength']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['DependenceNonUniformityNormalized', 'Maximum3DDiameter', 'SmallAreaEmphasis', 'Me
an', 'LargeDependenceEmphasis', 'DependenceEntropy', 'Maximum', 'GrayLevelVariance2', 'Minimum', 'Mea
nAbsoluteDeviation', 'MinorAxisLength', 'LargeAreaEmphasis', 'JointAverage', 'SmallAreaHighGrayLevelE
mphasis', 'Energy', 'LongRunHighGrayLevelEmphasis', 'JointEntropy', 'ENOLISMO', 'SmallDependenceLowGr
ayLevelEmphasis']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['Maximum', 'SmallDependenceLowGrayLevelEmphasis', 'Minimum', 'SmallAreaEmphasis',
'JointEntropy', 'MeanAbsoluteDeviation', 'DependenceEntropy', 'LongRunHighGrayLevelEmphasis', 'JointA
verage', 'GrayLevelVariance2', 'Energy', 'SmallAreaHighGrayLevelEmphasis']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['SmallAreaEmphasis', 'JointEntropy', 'Energy', 'SmallAreaHighGrayLevelEmphasis',
'JointAverage']
            Time taken for regular XGBoost feature selection = 0 seconds
```

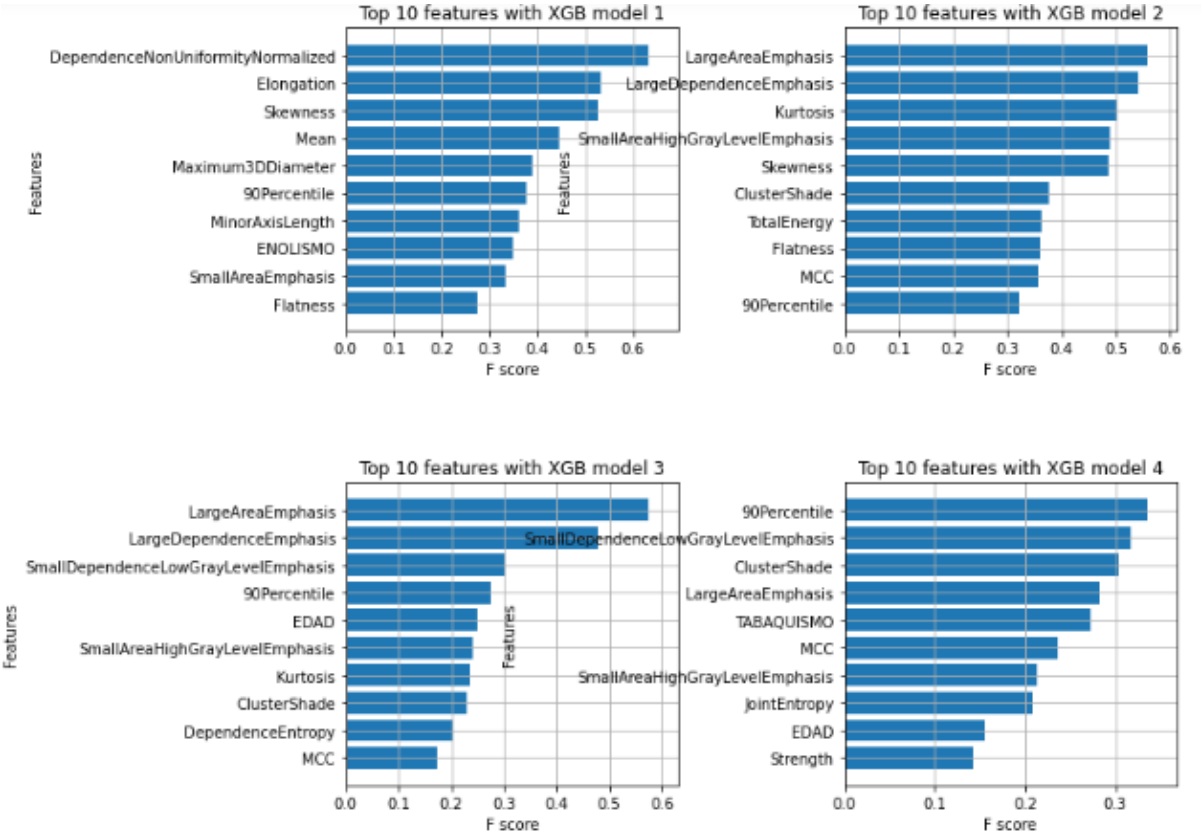


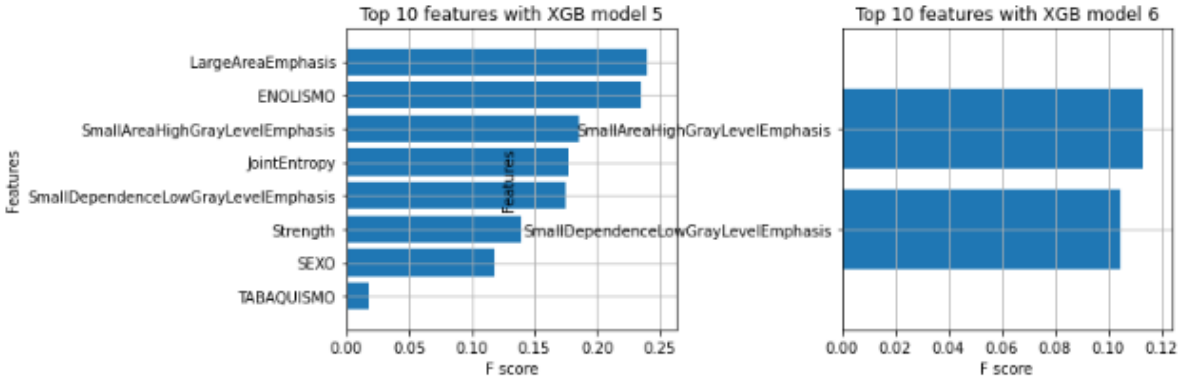



```
    Completed XGBoost feature selection in 1 seconds
################################################################################################
#####          F E A T U R E   S E L E C T I O N   C O M P L E T E D            #######
################################################################################################
Selected 33 important features. Too many to print...
Total Time taken for featurewiz selection = 3 seconds
Output contains a list of 33 important features and a train dataframe
#########################################################################################################
############       F A S T   F E A T U R E  E N G G   A N D   S E L E C T I O N ! ########
# Be judicious with featurewiz. Don't use it to create too many un-interpretable features! #
#########################################################################################################
featurewiz has selected 0.8 as the correlation limit. Change this limit to fit your needs...
Skipping feature engineering since no feature_engg input...
Skipping category encoding since no category encoders specified in input...
#### Single_Label Multi_Classification problem ####
    Loaded train data. Shape = (59, 34)
    Some column names had special characters which were removed...
#### Single_Label Multi_Classification problem ####
No test data filename given...
################################################################################################
######################## C L A S S I F Y I N G  V A R I A B L E S  ###################
################################################################################################
       No variables were removed since no ID or low-information variables found in data set
No GPU active on this device
    Tuning XGBoost using CPU hyper-parameters. This will take time...
Removing 0 columns from further processing since ID or low information variables
    After removing redundant variables from further processing, features left = 33
No interactions created for categorical vars since feature engg does not specify it
################################################################################################
#####  Searching for Uncorrelated List Of Variables (SULOV) in 33 features ############
################################################################################################
    there are no null values in dataset...
    Removing (1) highly correlated variables:
    ['JointAverage']
```



```
Time taken for SULOV method = 0 seconds
    Adding 0 categorical variables to reduced numeric variables  of 32
Finally 32 vars selected after SULOV
Converting all features to numeric before sending to XGBoost...
###############################################################################################
#####    R E C U R S I V E   X G B O O S T : F E A T U R E   S E L E C T I O N  #######
###############################################################################################
    using regular XGBoost
    Taking top 8 features per iteration...
    XGBoost version using 1.7.5 as tree method: hist
    Number of booster rounds = 100
        Selected: ['DependenceNonUniformityNormalized', 'Elongation', 'SmallAreaEmphasis', 'Maximum3D
Diameter', 'Mean', 'MeanAbsoluteDeviation', 'MinorAxisLength', 'Flatness']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['LargeAreaEmphasis', 'Flatness', 'MeanAbsoluteDeviation', 'GrayLevelVariance2', 'D
ependenceEntropy', 'Sphericity', 'LargeDependenceEmphasis', 'Skewness']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['LargeDependenceEmphasis', 'LargeAreaEmphasis', 'ClusterShade', 'SmallDependenceLo
wGrayLevelEmphasis', 'DependenceEntropy', 'Minimum', 'Kurtosis', '90Percentile']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['LargeAreaEmphasis', 'SmallDependenceLowGrayLevelEmphasis', 'ClusterShade', '90Per
centile', 'JointEntropy', 'MCC', 'EDAD', 'Energy']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['LargeAreaEmphasis', 'JointEntropy', 'SmallDependenceLowGrayLevelEmphasis', 'Small
AreaHighGrayLevelEmphasis', 'Strength', 'SEXO', 'ENOLISMO', 'TABAQUISMO']
            Time taken for regular XGBoost feature selection = 0 seconds
        Selected: ['SmallDependenceLowGrayLevelEmphasis', 'SmallAreaHighGrayLevelEmphasis']
            Time taken for regular XGBoost feature selection = 0 seconds
```






```
    Completed XGBoost feature selection in 1 seconds
################################################################################################
#####          F E A T U R E   S E L E C T I O N   C O M P L E T E D            #######
################################################################################################
Selected 28 important features:
['DependenceNonUniformityNormalized', 'Elongation', 'SmallAreaEmphasis', 'Maximum3DDiameter', 'Mean',
'MeanAbsoluteDeviation', 'MinorAxisLength', 'Flatness', 'LargeAreaEmphasis', 'GrayLevelVariance2', 'D
ependenceEntropy', 'Sphericity', 'LargeDependenceEmphasis', 'Skewness', 'ClusterShade', 'SmallDepende
nceLowGrayLevelEmphasis', 'Minimum', 'Kurtosis', '90Percentile', 'JointEntropy', 'MCC', 'EDAD', 'Ener
gy', 'SmallAreaHighGrayLevelEmphasis', 'Strength', 'SEXO', 'ENOLISMO', 'TABAQUISMO']
Total Time taken for featurewiz selection = 2 seconds
Output contains a list of 28 important features and a train dataframe
```

```
train2.head()
```

NumTAC	DependenceNonUniformityNormalized	Elongation	SmallAreaEmphasis	Maximum3DDiameter	Mean	MeanAbsolut
1	0.180107	0.763887	0.664810	23.805090	-203.250000	
2	0.203488	0.789238	0.722510	24.940607	-178.400943	
3	0.195262	0.785386	0.704023	16.160783	-93.074689	
5	0.087984	0.701980	0.715111	35.108607	-68.407331	
6	0.052678	0.795386	0.642570	34.187500	8.256217	

5 rows × 29 columns

8.1.3.2 Select K Best

Reduim a un màxim de 10 característiques radiòmiques

Per a variables clíniques agafem totes les que surten del featurewiz

```
# Inicializar SelectKBest con la función de prueba f_classif (ANOVA F-value)
selector = SelectKBest(score_func=f_classif, k=10)  # Seleccionar las  mejores características

# Aplicar SelectKBest a los datos
X_nuevo = selector.fit_transform(train2, y)

# Obtener las características seleccionadas
selected_features = train2.columns[selector.get_support()].tolist()
print(selected_features)
# Seleccionar solo las variables 10 mejores
selected_data = train2[selected_features]
```

```
['DependenceNonUniformityNormalized', 'Elongation', 'Maximum3DDiameter', 'Mean', 'MinorAxisLength',
'DependenceEntropy', 'LargeDependenceEmphasis', '90Percentile', 'Strength', 'RESPUESTA']
```

8.1.4 Resultats després del pre-processat

```
#split data into feature and target
X_new = selected_data.drop([ 'RESPUESTA' ],axis= 1 )
y_new = selected_data.RESPUESTA.values

#Estandarice las características usando StandardScaler de scikit-learn.
X_scaled = StandardScaler().fit_transform(X_new)

# Inicializar el validador cruzado (Stratified K-Fold)
n_splits = 5
skf = StratifiedKFold(n_splits=n_splits, shuffle=True, random_state=42)

# Inicializar una lista para almacenar los resultados
accuracy_scores = []
all_y_true = []
all_y_pred = []
all_y_prob = []
# Inicializar una lista para almacenar las matrices de confusión
confusion_matrices = []

# Ciclo de validación cruzada
for train_index, test_index in skf.split(X_scaled, y_new):
    X_train, X_test = X_scaled[train_index], X_scaled[test_index]
    y_train, y_test = y_new[train_index], y_new[test_index]

    # Crear el modelo de Random Forest Classifier
    rf_classifier = RandomForestClassifier(n_estimators = 100, criterion = 'gini', max_depth = None,
                                    min_samples_split = 2, min_samples_leaf = 1,
                                    min_weight_fraction_leaf = 0.0,max_features = 'sqrt',
                                    max_leaf_nodes = None, min_impurity_decrease = 0.0,
                                    bootstrap = True, oob_score = True, n_jobs = None,
                                    random_state = None, verbose = 0,
                                    warm_start = False, class_weight = None, ccp_alpha = 0.0)
    # Entrenar el modelo en el conjunto de entrenamiento
    rf_classifier.fit(X_train, y_train)

    # Realizar predicciones en el conjunto de prueba
    y_pred = rf_classifier.predict(X_test)
    # Obtener las probabilidades de clase para la clase positiva (1)
    y_prob = rf_classifier.predict_proba(X_test)
      # Almacenar las etiquetas reales y las predicciones
    all_y_true.extend(y_test)
    all_y_pred.extend(y_pred)
    all_y_prob.extend(y_prob)

    # Calcular la exactitud y almacenarla en la lista
    accuracy = accuracy_score(all_y_true, all_y_pred)
    accuracy_scores.append(accuracy)
```

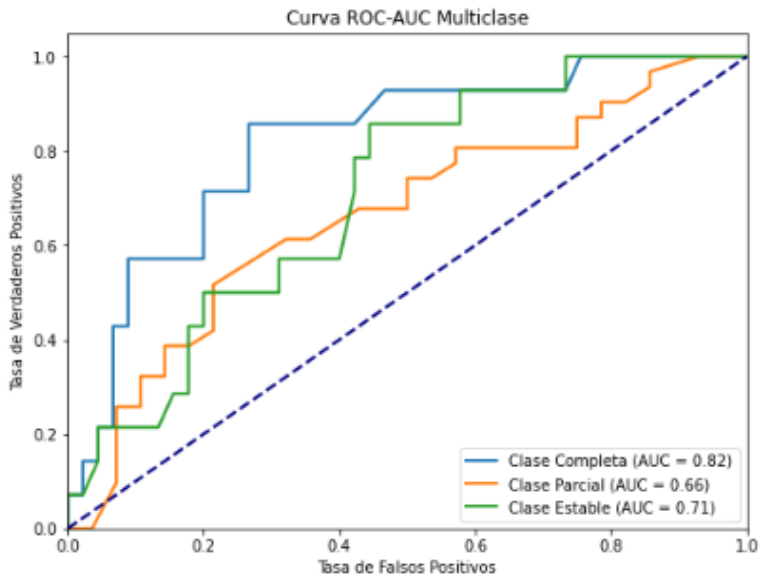
```
# Calcular la exactitud promedio y desviación estándar de las puntuaciones
average_accuracy = np.mean(accuracy_scores)
std_accuracy = np.std(accuracy_scores)

print("Exactitud promedio:", average_accuracy)
print("Desviación estándar de la exactitud:", std_accuracy)
```

```
Exactitud promedio: 0.6089218455743879
Desviación estándar de la exactitud: 0.02039829271845976
```

8.1.5 Mètriques d'avaluació

8.1.5.1 Corba AUC-ROC

```
# Inicializar listas para almacenar los resultados de ROC y AUC para cada clase
roc_curves = []
auc_scores = []

# Calcular la curva ROC y el AUC para cada clase
for i in range(len(all_y_prob[0])):
    fpr, tpr, _ = roc_curve([1 if label == i else 0 for label in all_y_true],
                            [pred[i] for pred in all_y_prob])
    roc_curves.append((fpr, tpr))
    auc_scores.append(roc_auc_score([1 if label == i else 0 for label in all_y_true],
                                    [pred[i] for pred in all_y_prob]))

# Graficar la curva ROC para cada clase
plt.figure(figsize=(8, 6))
# Etiquetas personalizadas para las clases
clase = ['Completa', 'Parcial', 'Estable']
for i, (fpr, tpr) in enumerate(roc_curves):
    plt.plot(fpr, tpr, lw=2, label=f'Clase {clase[i]} (AUC = {auc_scores[i]:0.2f})')

# Graficar la línea diagonal (aleatoria)
plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--')

plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.05])
plt.xlabel('Tasa de Falsos Positivos')
plt.ylabel('Tasa de Verdaderos Positivos')
plt.title('Curva ROC-AUC Multiclase')
plt.legend(loc="lower right")
plt.show()

# Imprimir el AUC macro (promedio de los AUC de todas las clases)
print(f'AUC Macro: {np.mean(auc_scores):0.2f}')
```




```
AUC Macro: 0.73
```

8.1.5.2 Matriu de confusió

```
# Calcular la matriz de confusión y almacenarla en la lista
from sklearn.metrics import confusion_matrix
from matplotlib.ticker import FixedLocator

confusion = confusion_matrix(all_y_true, all_y_pred)
confusion_matrices.append(confusion)

# Calcular la matriz de confusión promedio
cm = np.mean(confusion_matrices, axis=0)

# Obtener los valores verdaderos positivos (VP) y falsos positivos (FP) para cada variable
VP = np.diag(cm)
FP = np.sum(cm, axis=0) - VP
# Calcular el porcentaje de predicción para cada variable
percentage_pred = VP / (VP + FP)


# Obtener la diagonal principal de la matriz de confusión
diagonal = np.diagonal(cm)

# Calcular el porcentaje de acierto
accuracy = np.sum(diagonal) / np.sum(cm)

# Crear un mapa de calor de la matriz de confusión
sns.heatmap(cm, annot=True, cmap='Blues')

# Etiquetas personalizadas para las clases
class_labels = ['Completa', 'Parcial', 'Estable']

# Personalizar las ubicaciones y etiquetas en el eje x
num_classes = len(class_labels)
x_positions = np.arange(num_classes)
x_labels = class_labels

# Mover las etiquetas hacia la derecha (ajusta el valor según sea necesario)
x_positions_shifted = x_positions + 0.5
y_positions_shifted = x_positions + 0.5

# Configurar las ubicaciones de las etiquetas en el eje x
plt.gca().xaxis.set_major_locator(FixedLocator(x_positions_shifted))
plt.gca().set_xticklabels(x_labels, rotation=0, ha='right')
# Configurar las ubicaciones de las etiquetas en el eje x
plt.gca().yaxis.set_major_locator(FixedLocator(y_positions_shifted))
plt.gca().set_yticklabels(x_labels, rotation=0, ha='right')
```

```
plt.title ('Matriu de confusió (Característiques radiòmiques)')
plt.xlabel('Etiquetas predichas', fontweight='bold')
plt.ylabel('Etiquetas reales', fontweight='bold')
for i, label in enumerate(class_labels):
    plt.text(i + 0.5, i + 0.5, f'{percentage_pred[i]:.2%}', fontsize = 20, ha='center',
             va='center', color='red')

plt.show()
```