

Projecte fi de grau

Estudi: Grau en Enginyeria Informàtica

Títol: Aprenentatge automàtic per a la generació de plans de tractament de tumors amb radioteràpia

Document: Resum

Alumne: Joan Passarrius Pedrosa

Tutor: Beatriz López Ibáñez  
Departament: Enginyeria Elèctrica, Electrònica i Automàtica  
Àrea: Enginyeria de Sistemes i Automàtica

Tutor extern: Daniel Lambisto Castro  
Departament: Institut Català d'Oncologia (ICO)  
Àrea: Servei de Física Mèdica i Protecció Radiològica

Convocatòria (mes/any): Febrer 2022

PROJECTE FI DE GRAU

---

# Aprentatge automàtic per a la generació de plans de tractament de tumors amb radioteràpia

---

*Autor:*

Joan PASSARRIUS PEDROSA

Febrer 2022

Grau en Enginyeria Informàtica

*Tutors:*

Beatriz LÓPEZ IBÁÑEZ

Daniel LAMBISTO CASTRO

# Introducció

---

La radioteràpia consisteix en l'ús de radiacions ionitzants per al tractament de determinats tipus de tumors. És una teràpia oncològica, com pot ser la cirurgia o la quimioteràpia, que elimina, amb les radiacions, les cèl·lules canceroses. El problema d'aquesta tècnica és que aconseguir un tractament on la radiació només afecti la regió ocupada pel tumor i que, a més, aquest tractament sigui el més efectiu possible, és molt complicat. Això es deu a que el procés de generar un pla de tractament per cada pacient és bastant complex donat el nombre de factors que s'han de tenir en compte a l'hora de realitzar els càlculs per acabar generant cada model, i el temps que requereix la generació de cadascun d'ells.

## 1.1 Motivació

El motiu principal que m'ha portat a escollir aquest projecte és el fet de poder posar en pràctica els conceptes adquirits durant la carrera i poder aprofundir més en els coneixements relacionats amb la Intel·ligència Artificial.

Un altre aspecte molt rellevant que m'ha fet decidir per aquest estudi ha estat el fet que el projecte està relacionat amb l'àmbit de la medicina. Això otorga un extra de motivació, ja que a part d'ampliar els coneixements propis, també serveix per ajudar en l'avenç i millora de procediments mèdics, que al final repercutiran en augmentar tant l'eficàcia com la velocitat de generació de tractaments podent oferir els millors serveis a persones que pateixen qualsevol tipus de tumor.

Així doncs he considerat que el projecte d'en Daniel (tutor extern de l'ICO) és perfecte, ja que la idea que proposa és molt atractiva en quant a l'aspecte de recerca i innovació, i s'ajusta molt als objectius personals que m'agradaria assolir.

## 1.2 Propòsit

El propòsit principal d'aquest projecte consisteix en estudiar la relació entre les diferents mètriques que s'utilitzen per generar els tractaments de radioteràpia utilitzant tècniques de Machine Learning per tal d'aconseguir automatitzar el procés de generació de plans per al tractament de tumors amb radioteràpia, a més d'aconseguir desenvolupar un model predictiu capaç de predir tractaments òptims per a cada pacient reduint al mínim l'afectació de la dosi en els teixits sans i, per tant, reduir-ne la toxicitat i els efectes adversos.

L'abast que s'ha establert comprèn l'estudi, anàlisi i disseny d'un sistema predictiu programat en Python capaç de predir plans de tractament on es minimitzi l'error utilitzant tècniques de Machine Learning.

## 1.3 Objectius

Els objectius d'aquest P/TFG són els següents:

- Estudiar el comportament de les diferents mètriques utilitzades en un pla de tractament obtingut a través d'un *planificador* i trobar quines són les que aporten més informació, i per tant, tenen una major relació a l'hora de decidir si un tractament passa els índexs de validesa per tal de poder-lo utilitzar per a tractar un pacient. Totes les mètriques que s'utilitzen per elaborar un pla de tractament són les recomanades per la Societat Catalana de Física Mèdica, però no totes aquestes mètriques tenen per què ser rellevants i donar-nos la mateixa informació. En concret s'estudiarà la utilització de diferents mètodes de regressió per a seleccionar les millors mètriques.
- Fer l'anàlisi, el disseny i la implementació d'un sistema intel·ligent que permeti obtenir prediccions sobre quins són els tractaments que passen l'índex d'acceptació per poder ser utilitzats en un pacient rebent d'entrada les mètriques seleccionades.

## CAPÍTOL 2

# Requisits

---

S'ha determinat que el sistema resultant ha de complir amb els requeriments següents:

- Donades les dades de múltiples tractaments de radioteràpia, obtenir la informació referent al seu comportament (lineal o no lineal) i detectar les mètriques més rellevants.
- Classificar un tractament de radioteràpia com vàlid o invàlid. Es tracta d'una classificació binària.

## CAPÍTOL 3

# Resultats

---

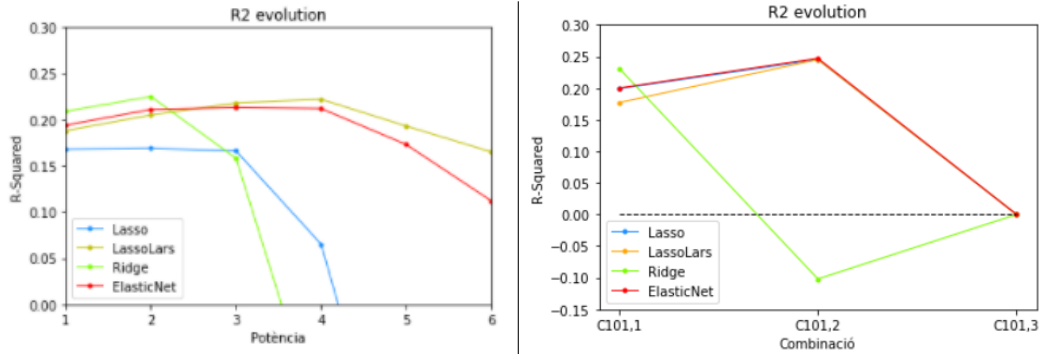
En aquest apartat es mostren de forma gràfica els resultats obtinguts tant al realitzar l'anàlisi de les mètriques utilitzant regressions, com en realitzar la classificació dels tractaments utilitzant diferents algorismes d'aprenentatge automàtic.

S'han realitzat les proves amb dos conjunts de mètriques diferents. Un d'ells conté totes les mètriques, mentre que l'altre està format per un subconjunt seleccionat a partir de la seva influència estadística.

S'han aplicat diferents processos als conjunts de mètriques per tal de generar els escenaris on s'han realitzat les proves. Amb el primer procés, s'han generat les potències de fins a grau  $p=6$  de les mètriques per tal d'avaluar la dependència de les variables amb elles mateixes, mentre que amb el segon s'han generat combinacions amb repeticions de les variables per determinar l'existència de dependències creuades. D'aquesta forma, per cada conjunt de mètriques s'han obtingut 9 escenaris diferents, 6 amb el primer procés i 3 amb el segon.

### 3.1 Regressions

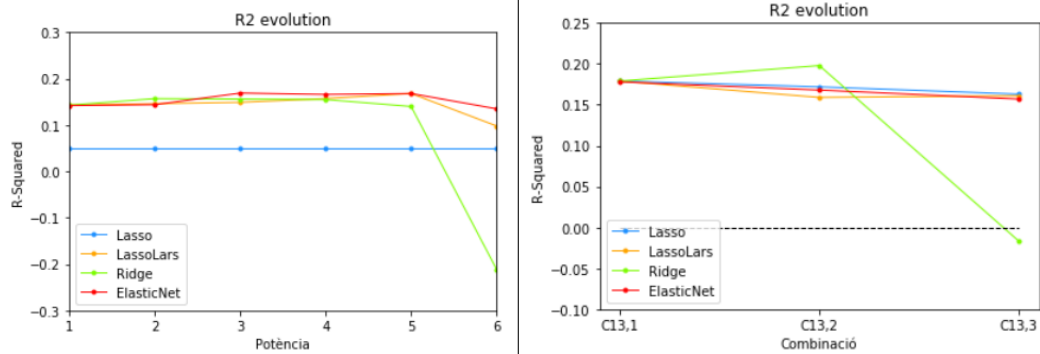
Els resultats obtinguts amb les regressions han estat el següents:



(a) Gràfic amb l'evolució dels coeficients de determinació segons la potència i el model de regressió.

(b) Gràfic amb l'evolució dels coeficients de determinació segons la combinació de variables i el model de regressió.

Figura 3.1: Resultats obtinguts amb models de regressió lineal provats amb totes les mètriques dels tractament.



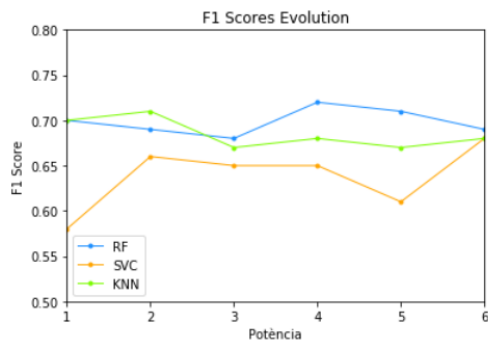
(a) Gràfic amb l'evolució dels coeficients de determinació segons la potència i el model de regressió.

(b) Gràfic amb l'evolució dels coeficients de determinació segons la combinació de variables i el model de regressió.

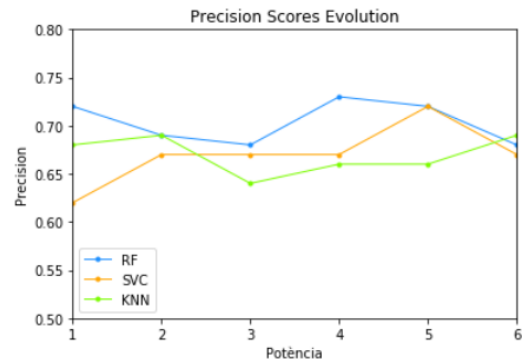
Figura 3.2: Resultats obtinguts amb models de regressió lineal provats amb un subconjunt de les mètriques dels tractament.

### 3.2 Classificadors

Els resultats obtinguts amb els classificadors han estat el següents:

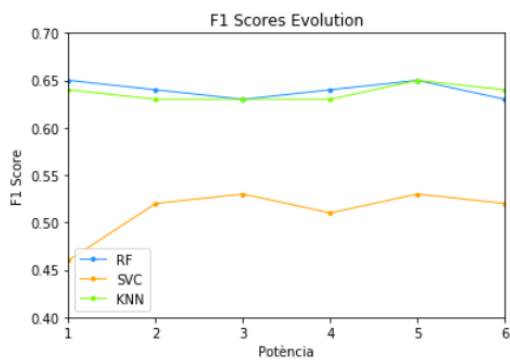


(a) Representació gràfica de l'evolució de les  $F_1$  Score de cada classificador segons la potència de les mètriques.

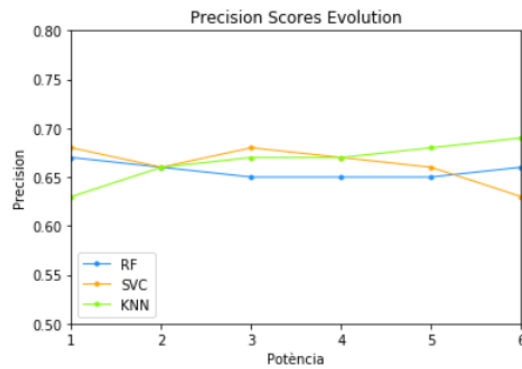


(b) Representació gràfica de l'evolució de les precisions de cada classificador segons la potència de les mètriques.

Figura 3.3: Resultats dels classificadors utilitzant totes les mètriques dels tractaments.



(a) Representació gràfica de l'evolució de les  $F_1$  Score de cada classificador segons la potència de les mètriques.



(b) Representació gràfica de l'evolució de les precisions de cada classificador segons la potència de les mètriques.

Figura 3.4: Resultats dels classificadors utilitzant només un subconjunt de les mètriques dels tractaments.

## CAPÍTOL 4

# Conclusions

---

El caràcter d'aquest treball està encarat de forma explícita a la recerca i té com objectiu determinar si és possible fer una predicció de la viabilitat dels tractaments de radioteràpia a través de la informació que s'obté de les diferents mètriques que defineixen aquests tractaments.

Amb aquest enfocament es van definir dues etapes del projecte, la primera consistia en l'estudi de les mètriques i, l'altra consistia en el desenvolupament d'un model predictiu que fos capaç de predir amb les mètriques seleccionades en l'estudi de l'etapa anterior quins tractaments són els que superen els índexs de viabilitat establerts.

A trets generals podem dir que els objectius marcats s'han assolit:

- S'han estudiat les mètriques i s'han extret conclusions interessants que marquen un nou full de ruta per seguir la investigació. La principal conclusió és que la correlació entre la viabilitat del tractament i les mètriques de complexitat analitzades indica que, d'existir relació entre elles, aquesta no és una relació lineal, per tant s'hauria de seguir la investigació utilitzant regressions no lineals o altres mètodes.

A més, s'ha descobert una característica important de les dades que és la multicolinealitat i s'ha determinat que intentar-la solventar pot comportar en la majoria dels casos la pèrdua d'informació rellevant. Amb això s'ha establert que per futurs estudis, és imprescindible fer ús de totes les mètriques.

- Utilitzant els mètodes de Machine Learning inclosos dins la llibreria de Scikitlearn de Python s'ha aconseguit obtenir un model que és capaç de predir en un 72% d'efectivitat quins són els tractaments que passen els índexs de viabilitat per poder ser utilitzats en pacients sense necessitat de fer proves amb maniquí. Tot i no ser un bon resultat donada la finalitat del predictor, ja que l'existència de falsos positius n'impossibiliten el seu ús en casos reals, s'ha pogut demostrar que no és possible fer una predicció de la viabilitat dels tractaments de radioteràpia utilitzant els mètodes de boscos aleatoris, màquines de suport vectorial, o KNN.