Objectives

The goal of this project is to develop a new deep learning model to analyze the Head and Neck Cancer dataset, in combination with public databases, to improve the prediction rate of patients' survival compared to models built on traditional radiomic features.

We would like to investigate and compare the performance of the deep learning-based method to the more conventional methods, such as hand-engineered radiomic features. The tumor *volume* feature is usually often used in clinic as a prognostic feature. The C-index of the volume for the set of patients at hand is around 0.65. In this case the objective would be to get better results by using a 3D convolutional neural network.

Since using this type of convolutional neural network is hard, and it can take some time to train and develop a model, this becomes and objective to accomplish too.

Scope

The first task will to learn and to understand how Neural Networks and specifically how Convolutional Neural Networks Work. This way I will have a fully understanding of the background that all these methods use create models for survival prediction.

The next task will be to set up and to run the *DeepSurv* python package on a local computer. Once running, all the different parts of the package should be tested to see which ones are best suited to be reused to create a new Survival Prediction Model. Since the package is not prepared to have images as an input, an improvement will be to add the possibility to pass medical images to train the survival model.

Afterwards, a deep learning model will be created starting from zero but trying to use some ideas from other completed projects. This model, unlike *DeepSurv*, will not be using the Cox Proportional Hazards because, in this case, the approach will be to compare pairs and not to get the hazard function. Since the C-index is computed using pairs of patients, a Siamese neural network will be built. This type of network is suited for comparison tasks such as face recognition, but this time it will be comparing the pair's survival time.

Building the network will be done in two steps, the first one will be to create shallow network as a prototype. Afterwards, continuing from the previous network, a deeper model will be created to test the full potential of deep learning.

Required knowledge

To develop this project, it will require knowledge from different subjects. Since it's machine learning related it's heavily linked to the Machine Learning (APA) subject. Also, it's related to the Artificial Intelligence (IA) subject which was an introductory course in all the different types of AI.

To develop a machine learning model, R or Python are the most common programming languages to work with. In this case I will be using Python which I learned in the Programming Languages course (LP).

All this subjects are related to Computer Science which is the specialty I chose to develop my Final Grade Project.

CS adequacy

To develop this project, I will be using machine learning related skills. This is one of the most popular fields in Computer Science currently. The project touches all the different steps needed to develop a deep learning model and uses different skills acquired during my specialization. So, since it's closely related to the Machine Learning (APA) subject, it's adequate to be developed as a Final Degree Project.

Professional skills

The needed skills will be:

- Know how a Machine Learning project works. This skill should help me in how to better organize code and what are the different steps required to develop this type of projects to avoid losing too much time, I should have a fully understanding of this part.
- Know different types of Neural Networks and what are the best parameters for each one. I will need a medium knowledge in this skill since I only will be using one type of neural network, however this should help me in what's the best way to start training and which are the best initial parameters to start training a NN.
- Know how to use Convolutional Neural Networks and why are they useful and in which areas. The project uses a 3D CNN, so to know how to properly use them is very important, I would like to have a high level of understanding how it works and what's the best way to use them.
- Image processing skills. I will be dealing with 3D images so I should be able to know how to open an image file and how to work with the different image values, I will need a medium knowledge for this skill.
- Python programming. Almost all the project will be developed in Python so I should get really comfortable with this programming language since this should help me develop more efficient code and take less time to do so.