Draft version December 13, 2021
Typeset using LaTeX **modern** style in AASTeX63

RapidPlan Knowledge Based Modelling Algorithm Overview*

Jeffrey Morais[1]

[1]*Department of Physics, McGill University, Montreal*

1. INTRODUCTION

Given a prescribed dose of ionizing radiation, one would like to maximize the amount received by the organ developing tumours, denoted as the clinical target volume (CTV), and minimize the amount received by the ambient organs currently unaffected by the tumours, denoted as the organs at risk (OAR). Generally this is done by creating a plan in advance, contouring organs or structures in patient scans and considering the dose-volume histogram (DVH), and the geometric-based expected dose (GED). The DVH gives distributions for radiation levels for given volume of structures, whereas the GED considers the physical limitations of the method whereby the radiation is delivered by computing the distances between each structure and the target surface for a given radiation field geometry. Fogliata et al. [2014a] In most cases, the manual construction of a plan is both resource and time intensive, and the process would improve in efficiency if there were quicker methods to achieve very similar if not identical results. The solution in question is the RapidPlan algorithm, which uses machine learning to train models for outputting plans for a given patient and their corresponding set of structures.

2. MACHINE LEARNING

Varian's RapidPlan solution uses supervised machine learning to output the predicted DVH and GED estimates. The following section outlines the underlying mathematics behind this.

A learning algorithm requires an input which is known as the training data set. The set takes the form of a finite sequence $S = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, where $x_i \in X$ is a D-dimensional feature vector describing a set of structures or rather a data point, and $y_i \in Y$ is the label corresponding to it's attribute x_i. Next, the algorithm requires an output: a predictive function known as a hypothesis that acts as a probabilistic model . Deisenroth et al. [2020] This acts as the output of the learning algorithm and takes the form of a function $f : X \to Y$, which attempts to predict y for arbitrary

jeffrey.morais@mail.mcgill.ca

* Based on literature provided by Varian & cited medical physics papers.

given x that is not in the training set. Put more formally, we can denote the learning algorithm as a map M taking the form:

$$M : \bigcup_{n \in N} (X \times Y)^n \to Y^X. \tag{1}$$

On the left hand side, $\bigcup_{n \in N}(X \times Y)^n$, represents the unions of all possible combinations of inputs X and labels Y in N data sets, where $(X \times Y)^i$ represents a set of training data with i elements. The right hand side, Y^X, is known as the hypothesis space F which contains the set of all possible predicting functions that take in initial data and generate labels. Wolf [2020]

To construct the learning algorithm, there are some probabilistic assumptions that must be made. One of which is that the pairs (x_i, y_i) are treated as values corresponding to random variables (X_i, Y_i). The variables themselves are independently distributed according to a probability measure $dP(x, y) \in X \times Y$, which is described as a function defined on a set of events in a probability space, given that is satisfies standard measure properties. Roussas [2014] The probability measure takes possible labels and assignments them to a probability value.

The goal of the learning algorithm is to find an accurate hypothesis f with respect to a chosen loss function $L : X \times Y \to \mathbb{R}$, which measures how far off $f(x)$ is from the respective label y. The average loss is defined as risk and is given by the functional:

$$R[f] \coloneqq \int_{X \times Y} dP(x, y)\ L(y, f(x)), \tag{2}$$

where the minimization of the risk produces a better hypothesis. The probability measure dP is not a given, only the training data set S is, thus the learning algorithm is tasked with minimizing the risk without having to evaluate R explicitly.

Before concluding this section with a brief overview of an example of risk minimization, a general overview of the learning algorithm process is discussed. Training data is supplied into the learning algorithm, where the inputs and labels are known. The learning algorithm develops predictors which act as the model that takes in inputs and outputs labels. The model is compared to test data and the risk is computed based on how far off the model is. The algorithm then makes corrections based on the risk and produces another model and the process is repeated through iterations until an accurate model is obtained. Supervised learning means that the inputs and labels are both known, whereas more general cases would employ neural networks where the algorithm classifies both of these, only requiring a much larger amount of data sets.

Finally, take a look at an example of risk minimization: empirical risk minimization for linear functions. Although risk R is not computed directly, the average loss of a model can be computed directly as follows:

$$\hat{R[f]} \coloneqq \frac{1}{N} \sum_{j=1}^{N} L(y_j, f(x_j)) \tag{3}$$

The method of minimizing $\hat{R}$ is known as empirical risk minimization. If we take $X \times Y = \mathbb{R}^D \times \mathbb{R}$ and $F = \{f : X \to Y \mid \exists z \in \mathbb{R}^D : f(x) = \langle z, x \rangle\}$ as the set of linear functions, then the minimizer of the empirical risk with a quadratic loss is given by:

$$\hat{R[z]} := \frac{1}{N} \sum_{j=1}^{N} (y_j - \langle z, x_j \rangle)^2. \tag{4}$$

Here, $\langle z, x_j \rangle$ is the inner or scalar product between the two vectors z and x_j. Wolf [2020]

3. RAPIDPLAN ALGORITHM

RapidPlan is a knowledge based program (KBP) that computes DVH prediction models to prescribe optimal dose distributions for a given patient with associated structures and field geometry. Fogliata et al. [2019] The learning algorithm is accessed within the Eclipse treatment planning sytem (TPS), which uses dose prescriptions from existing inputted treatment plans, along with patient-specific anatomy to generate the precise models. The models predict DVH distributions for the OARs and CTVs along with optimization objectives to better produce the estimated DVH ranges. Tinoco et al. [2020] An example of DVH estimates an objectives can be seen in Fig. 1.

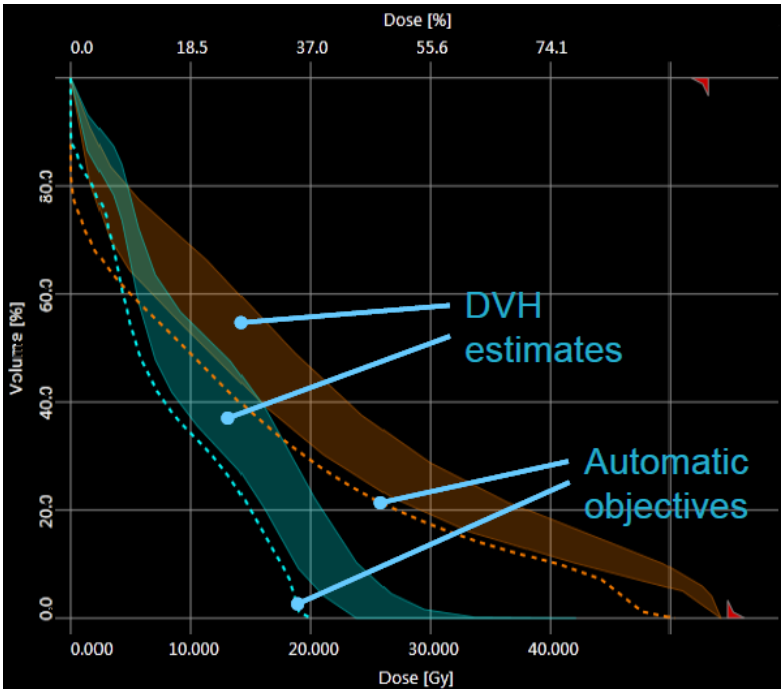

Figure 1. DVH distribution plot of volume percentage as a function of dose in units of Gy. Translucent regions represent DVH estimations computed by the model. Dotted lines represent DVH objectives. Phillips [2020]

Now, the steps for model generation of DVH distributions is discussed. First the algorithm requires a large set of existing patient training plans which each have a

structure set and an absolute dose for the CTV. Each of the plans are then associated to a collection of models whereby OARs, CTVs, and dose prescriptions can be defined. Once the principal components are defined, the KBP algorithm parametrizes structure sets and dose matrices, and defines dose-volume constraints on each structure. Next, geometric and dosimetric data is extracted from a patient databse to begin training the model. The training is based on principal component analysis methods. Fogliata et al. [2014b]

Once the model has been trained and is ready to be applied to patient data, patient information such as the structure sets, prescription dose and field geometry are inputed into the DVH estimation algorithm. After the DVH generation, automatic objectives are created based on the model in order to optimize the desired dose distribution. To optimize the current suggested plan, the algorithm calculates OAR volume partitions and GED cumulative volume histograms. Then using stepwise regression, coefficients are computed to create the most likely appropriate DVH, which can be seen in Fig. 2. Phillips [2020]

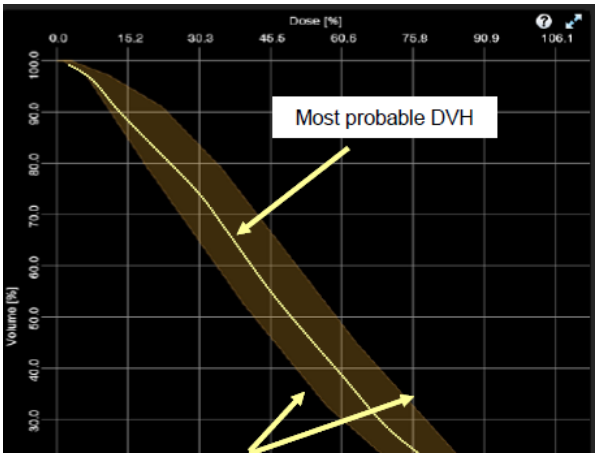

Figure 2. DVH distribution plot of volume percentage as a function of dose in units of Gy. Translucent regions represent DVH upper and lower bounds calculated for each OAR partition. Dotted line within the bounds represents the most probable DVH. Phillips [2020]

Once the most likely DVH is computed, the model is ready to be validated by the user before being applied to a patient plan.

REFERENCES

Marc Peter Deisenroth, A. Aldo Faisal, and Cheng Soon Ong. Mathematics for machine learning. 2020. https://doi.org/10.1017/9781108679930.

A. Fogliata, L. Cozzi, G. Reggiori, A. Stravato, F. Lobefalo, C. Franzese, D. Franceschini, S. Tomatis, and M. Scorsetti. Rapidplan knowledge based planning: iterative learning process and model ability to steer planning strategies. *Radiation Oncology*, 14(1), 2019. https://doi.org/10.1186/s13014-019-1403-0.

Antonella Fogliata, Po-Ming Wang, Francesca Belosi, Alessandro Clivio, Giorgia Nicolini, Eugenio Vanetti, and Luca Cozzi. Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer. *Radiation Oncology*, 9(1), 2014a. https://doi.org/10.1186/s13014-014-0236-0.

Antonella Fogliata, Po-Ming Wang, Francesca Belosi, Alessandro Clivio, Giorgia Nicolini, Eugenio Vanetti, and Luca Cozzi. Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer. *Radiation Oncology*, 9(1), 2014b. https://doi.org/10.1186/s13014-014-0236-0.

Helen Phillips. Rapidplan-algorithm. *Varian Oncology Medical Systmes*, (EC12.7-PPT-05-A), 2020. URL https://myvarian.com/.

George G. Roussas. Revised answers manual to an introduction to measure-theoretic probability. *An Introduction to Measure-Theoretic Probability*, 2014. https://doi.org/10.1016/b978-0-12-800042-7.00028-1.

Marisol Tinoco, Erika Waga, Kevin Tran, Hieu Vo, Jamie Baker, Rachel Hunter, Christine Peterson, Nicolette Taku, and Laurence Court. Rapidplan development of vmat plans for cervical cancer patients in low- and middle-income countries. *Medical Dosimetry*, 45(2):172–178, 2020. https://doi.org/10.1016/j.meddos.2019.10.002.

Michael M Wolf. Mathematical foundations of supervised learning. Jul 2020.