Evaluation of Machine Learning Algorithms for Treatment Planning Parameter Calculation

James Chow[1], Runqing Jiang[2] and Fredrick Ng[3]

[1] Princess Margaret Cancer Centre, Toronto, ON
[2] Grand River Regional Cancer Centre, Kitchener, ON
[3] Ryerson University, Toronto, ON

JULY 12–16 2020 VIRTUAL
JOINT AAPM | COMP MEETING
EASTERN TIME [GMT-4]

INTRODUCTION

To guarantee a high conformal dose deposited at the target, while simultaneously sparing the nearby organs-at-risk (OAR) in radiotherapy, a quality assurance (QA) program is needed to evaluate the dose plan to ensure it fulfils the minimum standard. Dose distribution index (DDI) is a relatively new dose-volume evaluation parameter proposed to help decision-making in treatment planning QA [1]. This index consolidates the dose coverage conformity and the homogeneity for the planning target volume (PTV), OAR, and remaining target-at-risk (RVR) into one value. This index can be determined solely from the prescribed dose and dose-volume histograms (DVHs) of the PTV, OAR, and RVR involved [2].

AIM

The aim of this study is to investigate how to predict the dose-volume parameter from dose-volume histograms in treatment planning QA using machine learning, and to evaluate the performance of different learning algorithms in the parameter prediction.

METHOD

DDIs of fifty prostate volumetric modulated arc therapy (VMAT) plans generated in the Grand River Hospital, Kitchener, ON, were calculated and compared to results predicted by machine learning. Different machine learning algorithms, namely, linear regression [3], tree regression (Fine, Medium and Coarse) [4], support vector machine (SVM) (Linear, Quadratic and Cubic) [5] and Gaussian process regression (GPR) (Square Experimental, Matern 5/2, Rational Quadratic and Exponential) [6] were used to predict the DDI based on the DVH in the plans. The learning, training and validation curve, root mean square error (RMSE), prediction speed and training time were determined for all algorithms in the performance evaluation.

To predict the DDI value based on DVHs from prostate VMAT plans using machine learning, a modified version of computer code generated using MATLAB's regression learner in the Statistical and Machine Learning toolbox App was used for this task. The workflow of machine learning is shown in Fig. 1.

RESULTS AND DISCUSSION

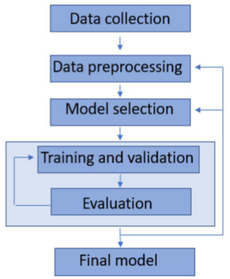

Figure 1: Flow chart for machine learning workflow: model selection, and data preprocessing may be repeated when desired performance is not reached.

Table 1. Summary of the performance of machine learning models. Table summarizing the RMSE, R-squared, prediction speed and training time of models created in the Regression Learning App available in the MATLAB's Machine Learning and Statistical Toolbox using five dose-volume points from each DVH with 4-fold cross validation. They are ordered from the best performance to worst.

Machine Learning Algorithm	RMSE	R-squared	Prediction Speed (Observation/s)	Training Time (s)
Square Exponential GPR	0.0038	0.99	4100	0.18
Matern 5/2 GPR	0.0038	0.99	3800	0.21
Rational Quadratic GPR	0.0038	0.99	2700	0.23
Linear Regression	0.0045	0.98	1700	0.37
Exponential GPR	0.0125	0.87	3900	0.18
Linear SVM	0.0123	0.87	4500	0.21
Quadratic SVM	0.0151	0.81	3400	0.13
Cubic SVM	0.0193	0.68	4700	0.11
Fine Tree	0.0218	0.60	4600	0.10
Medium Tree	0.0305	0.21	4600	0.42
Coarse Tree	0.0344	0 .00	5600	0.09

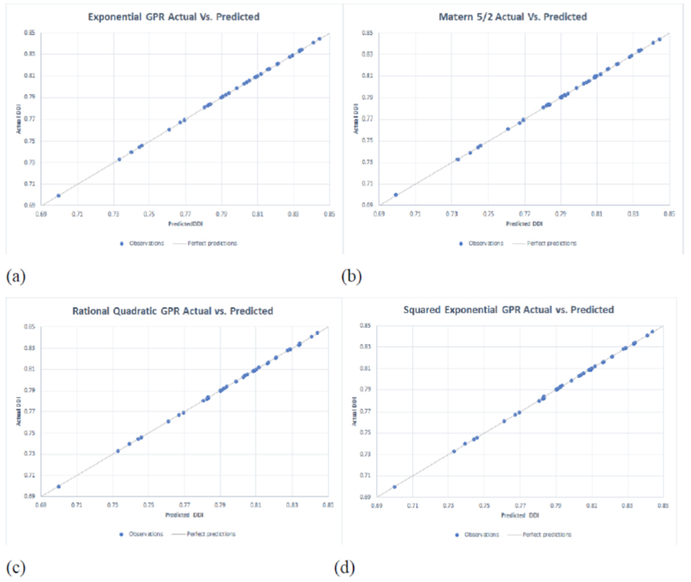

Figure 2: Predicted vs. actual plots of DDI values using the (a) exponential, (b) matern 5/2, (c) rational quadratic and (d) squared exponential Gaussian process regression algorithm in machine learning.

In this study, the GPR models (Fig. 2) were the most accurate among others. It was almost matched the calculated results with RMSE smaller than 0.004 (Table 1). The linear regression model was quite close in accuracy to the GPR. The downside is that the linear regression model took longer time (0.37 s vs. 0.18 s) to train, and had slower prediction speed (1,700 observation/s vs. 4,100 observation/s) when compared to the square exponential GPR. SVM models were performing reasonably well. They had the best prediction speed and training time out of other models that were accurate enough to reliably predict DDI. The linear SVM performed the best out of the SVMs.

CONCLUSIONS

Machine learning was used to predict the DDI from DVHs of prostate VMAT treatment plans using different models, namely, linear regression, tree regression, SVM, and GPR. It is found that all examined models were precise in their predictions of DDI, except the tree regression model. So it is possible to predict dose-volume parameter from the DVH data using machine learning. These models are not intended to replace already existing analytical calculations which are more interpretable. Instead, it provides a foundation of how machine learning can be used to determine dose-volume parameters where analytical solutions are not available.

ACKNOWLEDGEMENTS

We would like to thank Dr Runqing Jiang for her work on all prostate VMAT plans in the Grand River Hospital, Kitchener, ON.

REFERENCES

[1] Alfonso, JC, Herrero, MA, Nunez, L 2015 A dose-volume histogram based decision-support system for dosimetric comparison of radiotherapy treatment plans. *Radiat Oncol.* **10**(1) 263.
[2] Chow JCL, Jiang R, X Lu. Evaluation of plan optimizers in prostate VMAT using the dose distribution index. *J Radiother in Pract.* 2019;**18**(4):323-328.
[3] Matloff, N. 2017 Statistical Regression and Classification. Boca Raton, Florida, United States: CRC Press, pp.65-70.
[4] Loh, W. 2011 Classification and regression trees. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 1(1), pp.14-23.
[5] Suykens JA, Vandewalle J 1999 Least squares support vector machine classifiers. *Neural Proc lett.* **9**(3) 293-300.
[6] Seeger M 2004 Gaussian processes for machine learning. *Int J Neural Sys.* **14**(02) 69-106.

CONTACT INFORMATION

James Chow
Email: james.chow@rmp.uhn.ca
Runqing Jiang
Email: runqing.jiang@grhosp.on.ca
Fredrick Ng
Email: fnng@ryerson.ca