A Machine Learning approach for personalized dosimetry prediction in pediatric Nuclear Medicine applications

Vasilis Eleftheriadis[1], Georgios Savvidis[1], Panagiotis Papadimitroulas[1]

[1]BIOEMTECH, Mesogeion AV. 387, 15343, Athens, Greece

BIOEMTECH

INTRODUCTION

Patient-specific internal dosimetry for children is of great interest, due to the higher radiosensitivity that children experience in comparison to adults, as they may experience greater stochastic effects due to a higher percentage of replicating cells. The objective of the present study is the design, development and evaluation of an internal dosimetry prediction model with specific focus on pediatric patients, based on Machine Learning regression algorithms and AI techniques, which can accurately predict the specific absorbed dose rates (SADRs) in 31 organs, for five different radiopharmaceuticals.

MATERIALS AND METHODS

- The training and testing of the prediction models was performed using a simulated dosimetry database.
- GATE_v9.1 Monte Carlo toolkit[1] was used for the absorbed dose rate assessment at this study.
- 28 computational phantom (22 XCAT[2] and 6 ITIS[3]) with different anatomical characteristics and varying age, gender, mass and height were simulated.
- Five different radiopharmaceuticals (99mTc-MDP, 123I-mIBG, 131I-mIBG, 131I-Na and 153Sm-EDTMP), together with their specified activity distribution[4] were considered.
- Eight supervised ML regression algorithms were evaluated :
  - Least Squares Linear Regression
  - Ridge Regression
  - AdaBoost regressor
  - Gradient Boost regressor
  - XGBoost (eXtreme Gradient Boosting) regressor
  - Random Forest regressor
  - Decision Tree regressor
  - Support Vector Regressor (SVR)
- Combinations of 9 features, listed in Figure 1, were used as input for the regression algorithms.
- The Leave One Out Cross Validation (LOOCV)[5] method was used to train and validate the models.
- Models were trained on both single and multiple organ datasets (Figure2).
- Bayesian optimization[6] was performed on the 4 best performing algorithms (Gradient Boost, XGBoost, Random Forest, Decision Tree)
- Combinations of the 4 best performing algorithms were used to create Ensemble learning[7] models to boost predictive performance. (Figure 3)

RESULTS

- **99mTc-MDP**
  - Single organ XGBoost had the most accurate predictive ability with a Mean Absolute Percentage Error (MAPE) score of 8.85% on 27 organs, using the [0, 1, 2, 3, 4, 5, 6, 7, 8, 9] feature combination.
  - The ensemble 1-GB_3-RF_4-XGB_0-DT had the best MAPE score of 8.50% on 27 organs, using the [0, 1, 2, 3, 4, 5, 6, 7, 8, 9] feature combination.
- **123I-mIBG**
  - Multiple organ Gradient Boost had the best MAPE score of 8.19% on 27 organs, using the [0, 1, 2, 3, 4, 5, 6, 8] feature combination.
  - The ensemble 4-GB_0-RF_1-XGB_0-DT had the best MAPE score of 8.09% on 27 organs, using the [0, 1, 2, 3, 4, 5, 6, 8, 9] feature combination.
- **131I-mIBG**
  - Single organ Gradient Boost had the best MAPE score of 7.75% on 27 organs, using the [0, 1, 2, 3, 4, 5, 6, 7, 8, 9] feature combination.
  - The ensemble 4-GB_0-RF_1-XGB_0-DT had the best MAPE score of 7.69% on 27 organs, using the [0, 1, 2, 3, 4, 5, 6, 7, 8] feature combination.
- **131I-Na**
  - Multiple organ Gradient Boost had the best MAPE score of 8.09% on 22 organs, using the [0, 1, 2, 3, 4, 5, 6] feature combination.
  - The ensemble 4-GB_0-RF_1-XGB_0-DT had the best MAPE score of 7.69% on 22 organs, using the [0, 1, 2, 3, 4, 5, 6, 7, 8] feature combination.
- **153Sm-EDTMP**
  - Sinlge organ Gradient Boost had the best MAPE score of 7.96% on 26 organs, using the [0, 1, 2, 3, 4, 5, 6, 8] feature combination.
  - No ensemble managed to match Gradient Boost's performance in this radiopharmaceutical.

Organ	Time (h)	Age (y)	Gender	Weight (Kg)	Height (m)	BMI (kg/m²)	Torso to Top (cm)	Lung (cm)	Effective Diameter (cm)
0	1	2	3	4	5	6	7	8	9

Figure 1: The list of features used as input to the prediction models. An index was assigned to each of the features, for easier presentation of the input features' combinations.
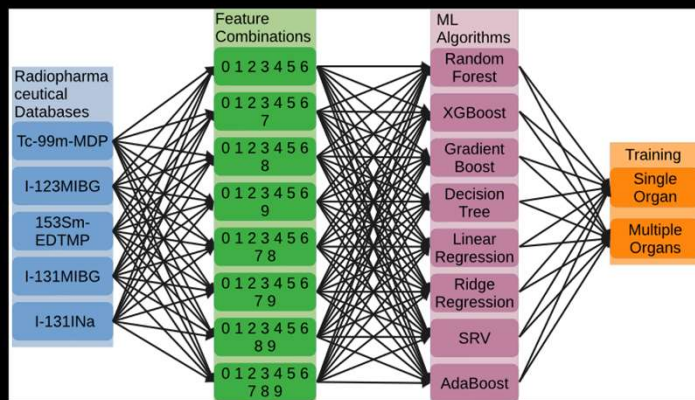

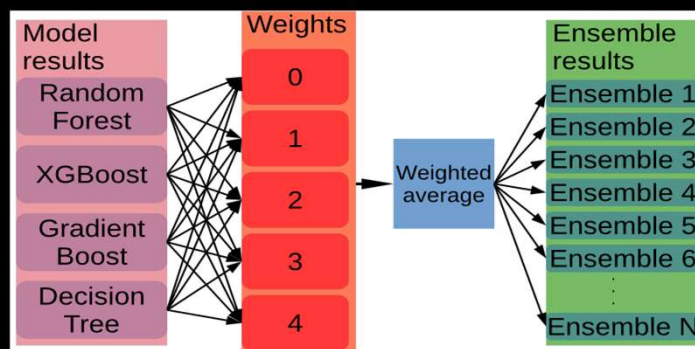Figure 2: Regression algorithms' training flowchart.

Figure 3 : Ensemble learning model creation flowchart.

DISCUSSION & CONCLUSIONS

- In this poster we present the methodology for the development of an internal dosimetry prediction model for pediatric patients based on Machine Learning and Artificial Intelligence techniques.
- The ML regression models where trained on a simulated dosimetry database.
- Using the right combination of input features, ML algorithms, Single or Multiple organs training and ensemble techniques that combine the best performing models we were able to predict the SADRs in 31 organs of pediatric patients, with a MAPE score of 8.50 or better, for five different radiopharmaceuticals.
- Although the results are within the acceptable threshold of error, training in larger and more diverse simulated cohorts should be considered a prerequisite before clinical use of the models.
- As a next steps, an online internal dosimetry prediction application for pediatric patients will be created, incorporating the prediction models designed during this study.

ACKNOWLEDGEMENTS

REFERENCES

[1] D. Sarrut et al., "A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications", 2014.
[2] WP Segars et al., "The development of a population of 4D pediatric XCAT phantoms for imaging research and optimization", 2015.
[3] A. Christ et al, "The Virtual Famliy-development of surface-based anatomical models of two adults and two children for dosimetric simulations", 2010.
[4] P. Papadimitroulas et al. "A personalized, Monte Carlo-based method for internal dosimetric evaluation of radiopharmaceuticals in children", 2018.
[5] C. Sammut et al. "Leave-One-Out Cross-Validation", 2011.
[6] Agnihotri et al. "Exploring Bayesian Optimization", 2020.
[7] T.G. Dietterich "Ensemble Methods in Machine Learning", 2000.