

## &lt;원저&gt;

## A Study on Predictive Modeling of I-131 Radioactivity Based on Machine Learning

Yeon-Wook You<sup>1)</sup>·Chung-Wun Lee<sup>1)</sup>·Jung-Soo Kim<sup>2)</sup><sup>1)</sup>Department of Nuclear Medicine and Research Institute, National Cancer Center<sup>2)</sup>Department of Radiological Science, Dongnam Health University

## 머신러닝 기반 고용량 I-131의 용량 예측 모델에 관한 연구

유연욱<sup>1)</sup>·이충운<sup>1)</sup>·김정수<sup>2)</sup><sup>1)</sup>국립암센터 핵의학과 · <sup>2)</sup>동남보건대학교 방사선학과

**Abstract** High-dose I-131 used for the treatment of thyroid cancer causes localized exposure among radiology technologists handling it. There is a delay between the calibration date and when the dose of I-131 is administered to a patient. Therefore, it is necessary to directly measure the radioactivity of the administered dose using a dose calibrator. In this study, we attempted to apply machine learning modeling to measured external dose rates from shielded I-131 in order to predict their radioactivity. External dose rates were measured at 1 m, 0.3 m, and 0.1 m distances from a shielded container with the I-131, with a total of 868 sets of measurements taken. For the modeling process, we utilized the hold-out method to partition the data with a 7:3 ratio (609 for the training set:259 for the test set). For the machine learning algorithms, we chose linear regression, decision tree, random forest and XGBoost. To evaluate the models, we calculated root mean square error (RMSE), mean square error (MSE), and mean absolute error (MAE) to evaluate accuracy and  $R^2$  to evaluate explanatory power. Evaluation results are as follows. Linear regression (RMSE 268.15, MSE 71901.87, MAE 231.68,  $R^2$  0.92), decision tree (RMSE 108.89, MSE 11856.92, MAE 19.24,  $R^2$  0.99), random forest (RMSE 8.89, MSE 79.10, MAE 6.55,  $R^2$  0.99), XGBoost (RMSE 10.21, MSE 104.22, MAE 7.68,  $R^2$  0.99). The random forest model achieved the highest predictive ability. Improving the model's performance in the future is expected to contribute to lowering exposure among radiology technologists.

**Key Words:** External Dose Rate, Machine Learning, Decision Tree, Random Forest, XGBoost

**중심 단어:** 공간선량률, 머신러닝, 의사결정트리, 랜덤포레스트, XGBoost

## 1. Introduction

The handling and administration of high-dose I-131 poses the risk of internal and external exposure for radiology technologists. The risk of internal exposure is less today than it has been in the past because of the increased stability of therapeutic capsule formulations of I-131 and the widespread adoption of the capsular form of I-131 for therapy [1, 2]. However, the risk of

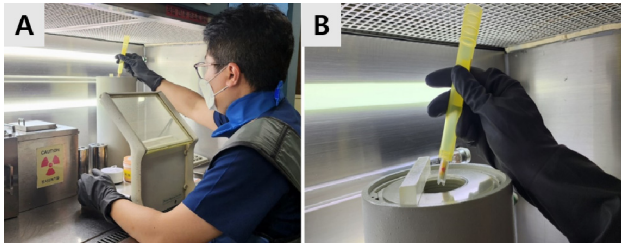
external exposure to localized areas such as the hands and eyes still exists because it is necessary to directly measure dose radioactivity with a dose calibrator when there is a difference between the calibration date and the date of the dose is administered to the patient. Radiology technologists should wear protective gear when measuring high-dose I-131 as shown in Fig. 1.

The equivalent dose can be calculated using Equation

Corresponding author: Jung-Soo Kim, Department of Radiological Science, Dongnam Health University, 50 Cheoncheon-ro 74-gil, Jangan-gu, Suwon-si, Gyeonggi-do, 16328, Republic of Korea / Tel: +82-31-249-6405 / E-mail: essencekim@dongnam.ac.kr

Received 26 January 2023; Revised 20 February 2023; Accepted 06 March 2023

Copyright ©2023 by The Korean Journal of Radiological Science and Technology



**Fig. 1.** Radiology technologists should wear protective gear (A) when measuring high-dose I-131 due to the risk of external exposure to localized areas such as the eyes and hands (B).

1, where  $W_R$  represents the radiation weighting factor and  $D_{T,R}$  is the absorbed dose averaged over the tissue or organ, T, of the incident radiation, R. For example, when working for 10 seconds to measurement 7400 MBq of I-131, the equivalent dose to a hand at a distance of 10 cm from the I-131 is 0.157 mSv (the gamma constant for I-131 is  $7.647E-5$  mSv/h per 1 MBq at 1 m).

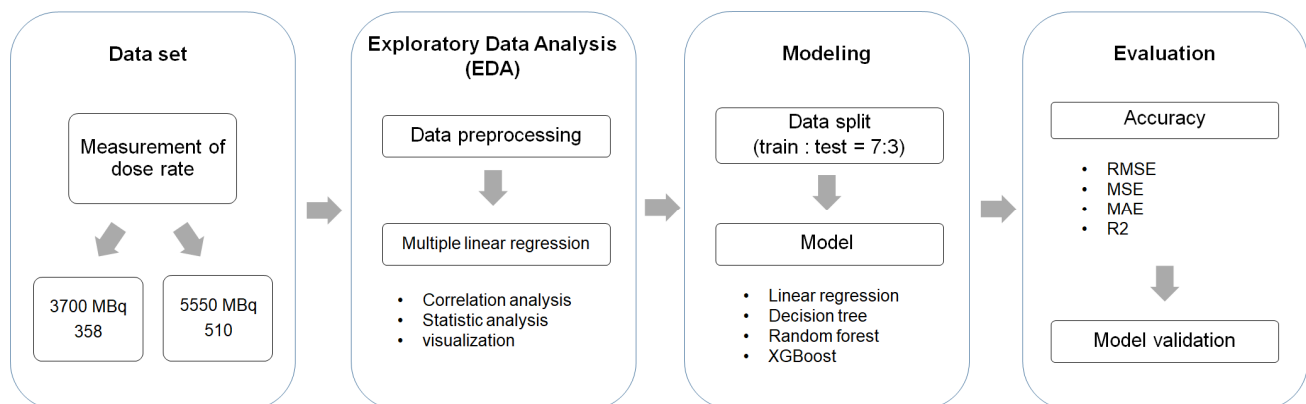
$$H_T = \sum_R W_R D_{T,R} \tag{1}$$

As the number of measurements increases, the exposure dose also increases. This causes great psychological burden on radiology technologists. In order to reduce exposure to radiation among workers in the long term, we conducted a study employing AI technology. Machine learning is a subtype of AI that uses algorithms through data analysis without being explicitly programmed [3–5]. Machine learning is generally divided into supervised learning and unsupervised learning. Supervised learning algorithms learn by using paired inputs and outputs presented to them by humans to find patterns so that

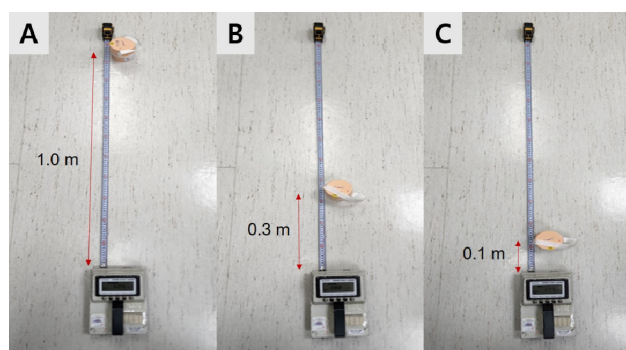
it can correctly predict output values for new data. In this study, we designed a supervised machine learning model by applying techniques used in big data analysis modeling, which is currently being used in the data science field. We aimed to train the model and predict the radioactivity of high dose I-131 based on the measured external dose rate of the I-131 in a shielded container.

## II. Materials and methods

A sequential depiction of our methodology is shown in Fig. 2. The external dose rates were measured at 1 m, 0.3 m and 0.1 m distances from I-131 sources in a shielded container obtained from suppliers, with the dose rate equal to the highest value measured over a 1-minute time span (Fig 3). Due to the characteristics of the survey meter for external dose rate, which measures in units of  $\mu$ Sv per hour, fluctuations in measured values can occur. In order to obtain a consistent measurement, we adopted the method of taking the highest value measured over a 1-minute time span. A total of 868 sets of measurements were acquired, 358 at 3700 MBq (100 mCi) and 510 at 5550 MBq (150 mCi). The actual measured activities are found to be significantly higher than the administered amount (up to 10% higher), due to the standard practice of preparing the dose with a margin, taking into account decay. Therefore, we explicitly distinguish between



**Fig. 2.** Methodology flowchart, we designed a supervised machine learning model by applying techniques used in big data analysis modeling, which is currently being used in the data science field.



**Fig. 3.** The external dose rates were measured at 1 m (A), 0,3 m (B) and 0,1 m (C) distances from I-131 sources in a shielded container, with the dose rate equal to the highest value measured over a 1-minute time span.

3700 MBq and 5550 MBq, not actual measured values. A RAM DA3-2000 Meter (Rotem Industries Ltd.) was used to measure the dose rates. We used only one type of instrument as we aimed to predict results using intuitive measurement values without correcting for correction factors.

## 1. Exploratory data analysis

We employed exploratory data analysis (EDA) which was devised by Tukey JW in 1977 [6]. Using EDA, the characteristics of the total data set were identified; data preprocessing, such as identifying outliers and missing values, was performed; and the correlation between explanatory variables and response variables was examined with a linear regression model. Finally, we visualized suitability of the data for modeling.

## 2. Machine learning

We used the hold-out method to generate the training and test sets for the training of the machine learning models. This is one of the simplest data resampling strategies: it randomly samples some data from the learning set for the test set, while the remaining data constitute the training set [7, 8]. In our study, the data set was split, with 70% of the data for the training set and 30% for the test set. We created two models which employed a simple algorithm and two models which employed an ensemble algorithm. All created supervised machine learning algorithms were implemented using

R programming language (version 4.2.0) in the Windows operating system environment. Linear regression is one of the oldest and most widely used correlational techniques. The goal of the method is to fit a straight line to a set of data points using a series of coefficients multiplied to each input, like a weighting function, and an intercept. Linear regression is easy to understand and quick to implement, even on larger data set. The downside of this method is that it is inherently linear and does not always fit real-world data [9]. A decision tree (DT) algorithm is used to divide learning activities where the tree is constructed by dividing the data set into smaller sets until each partition is clean and pure, where data classification depends on the type of data [10, 11]. The DT algorithm is one of the most effective learning algorithms due to its ability to handle all types of data, ease of comprehension and simplicity [12]. The party package in R was used with a maximum tree depth of 3 for visualization of DT (R Documentation: Package party version 1.3-10). In an ensemble algorithm, bagging or bootstrap aggregation (boosting) is a technique for reducing the variance of an estimated prediction function. Bagging seems to work especially well for high-variance, low-bias procedures, such as trees. For regression, we simply fit the same regression tree many times to bootstrap sampled versions of the training data, and average the results. Similar to bagging, boosting is a committee method, although, unlike bagging, the committee of weak learners evolves over time, and the members cast a weighted vote. Boosting appears to dominate bagging on most problems, and has become the preferred choice. We chose to use random forest (RF), which is a substantial modification of bagging that builds a large collection of de-correlated trees and then averages them. On many problems the performance of RF is very similar to boosting, and they are simpler to train and tune. As a consequence, RF is popular, and is implemented in a variety of packages [13]. Extreme gradient boost (XGBoost) is the fastest implementation of gradient boost, which is representative algorithm using boosting. XGBoost can harness all the processing power of modern multicore computers and is feasible to train on large data sets [14]. XGBoost was

chosen for comparison with RF.

### 3. Evaluation

To evaluate the predictive accuracy of the models, root mean square error (RMSE), mean square error (MSE), and mean absolute error (MAE) were calculated according to Equation 2–4, where  $y_i$  is the actual radioactivity of the I-131 dose and  $\hat{y}_i$  is the radioactivity predicted by the model.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{2}$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{3}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{4}$$

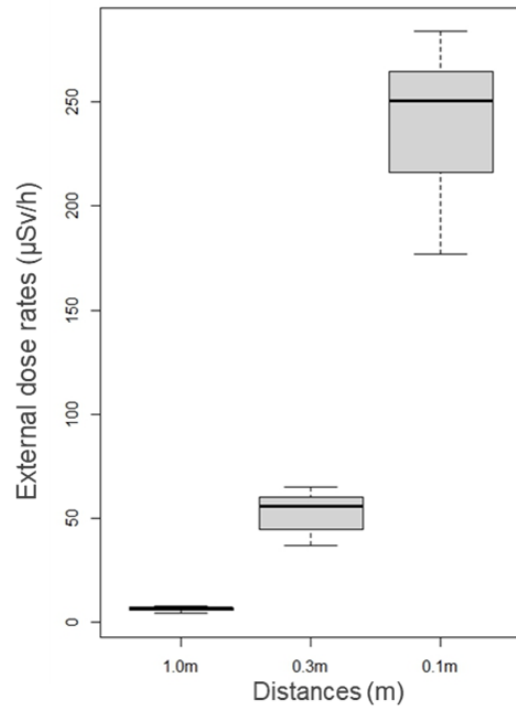
R squared ( $R^2$ ), the coefficient of determination, was also used to evaluate the models.  $R^2$  is calculated according to Equation 5, where  $\bar{y}_i$  is mean of the actual radioactivity of the I-131 dose ( $y_i$ ). The model-metrics package in R was used to print the evaluation results. To evaluate the two models with best performance, we validated them with 11 newly measured sets.

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2} \tag{5}$$

## III. Results

### 1. Exploratory data analysis

Fig. 4 is a boxplot of the total data set. There were no missing values and no outliers were detected. It can be seen that the interquartile range (IQR) of dose rate is the smallest at 1 m. This means that the dose rate does not fluctuate at increased distances from the shielded I-131, in accordance with the inverse-square law of distance. In the correlation analysis, the correlation coefficient between distance from the shielded I-131 and dose rate was high ( $\geq 0.9$ ), and the correlation coefficient between dose rate was also close to 1. We



**Fig. 4.** Boxplot of data set, x-axis is distance (m) from the shielded I-131, and y-axis is the external dose rates ( $\mu\text{Sv/h}$ ).

**Table 1.** Variance inflation factors for the total data set

Variable	VIF
1.0 m	22,742
0.3 m	45,458
0.1 m	51,145

\* Variance inflation factor

investigated multi collinearity by calculating the variance inflation factor (VIF) among the variables in the total data set. Since the VIF of the total data set was very high, we divided the data set into two groups, a 3700 MBq group and a 5550 MBq group, for statistical analysis (Table 1). The overall regression was statistically significant in both the 3700 MBq group ( $R^2=0.98$ ,  $F(3, 354)=5783.225$ ,  $P<0.001$ ) (Table 2) and the 5550 MBq group ( $R^2=0.96$ ,  $F(3, 506)=4494.678$ ,  $P<0.001$ ) (Table 3).

Fig. 5 is a visualization of the classification and regression trees (CART) algorithm implemented in the DT. The 0.3 m variable of the first node at the top implies that it is most closely related to the radioactivity of I-131. It can be seen that the data set was suitable for the model as it was split into 4 leaf nodes at 3700 MBq and 4 leaf nodes at 5550 MBq.

**Table 2.** Linear regression results for the 3,700 MBq group

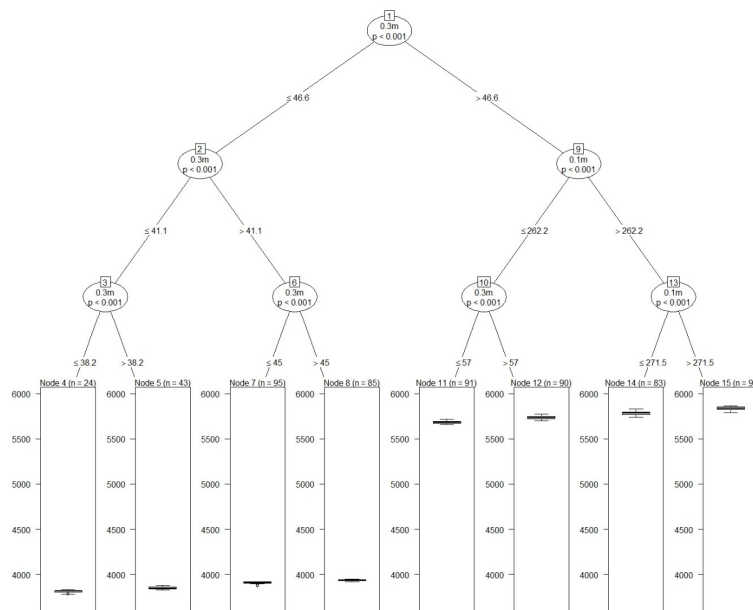
Dose 3,700 MBq	Coef <sup>‡</sup>	SE <sup>†</sup>	t <sup>‡</sup>	P	VIF <sup>§</sup>
Constant	3233.573	5.961	542.472	<0.001	
1.0 m	11.580	1.586	7.300	<0.001	3.127
0.3 m	8.612	0.337	25.558	<0.001	9.972
0.1 m	1.102	0.087	12.709	<0.001	10.755
Observations			358		
R <sup>2</sup>			0.980		
Adjusted R <sup>2</sup>			0.980		
Residual SE			6.136 (df = 354)		
F statistic			F = 5783.255, P < 0.001		

\* Coefficients, † Standard error, ‡ t-value, § Variance inflation factor

**Table 3.** Linear regression results for the 5,550 MBq group

Dose 5,550 MBq	Coef <sup>‡</sup>	SE <sup>†</sup>	t <sup>‡</sup>	P	VIF <sup>§</sup>
Constant	4357.549	12.907	337.620	<0.001	
1.0 m	38.947	4.206	9.260	<0.001	5.615
0.3 m	5.715	0.430	13.294	<0.001	6.123
0.1 m	3.035	0.130	23.416	<0.001	7.714
Observations			510		
R <sup>2</sup>			0.964		
Adjusted R <sup>2</sup>			0.964		
Residual SE			11.487 (df = 506)		
F statistic			F = 4494.678, P < 0.001		

\* Coefficients, † Standard error, ‡ t-value, § Variance inflation factor



**Fig. 5.** Visualization of the classification and regression trees (CART) algorithm implemented in the decision tree. The 0.3 m variable of the first node at the top implies that it is most closely related to the radioactivity of I-131. The data set was suitable for the model as it was split into 4 leaf nodes at 3700 MBq and 4 leaf nodes at 5550 MBq.

## 2. Evaluation of models

When evaluating the accuracy of the models, the smaller the RMSE, MSE and MAE, the better performance of the model, and the closer  $R^2$  was to 1, the higher explanatory power of the model. Random numbers used for random sampling were fixed using the seed function. The model with the best performance was the RF model, with a RMSE of 8.894, MSE of 79.098 and MAE of 6.546. On the other hand, as expected, the performance of the models employing a simple algorithm was not good. In particular, the MSE of the linear regression model was 71901.870 (Table 4).

## IV. Discussion

In this study, we created a data set consisting of external dose rates of high-dose I-131 measured at distances of 1 m, 0.3 m, 0.1 m, and used it to train supervised machine learning algorithms to predict the actual level of radioactivity of I-131 doses. The purpose of this study was not to compare and evaluate machine learning algorithms, but to focus on finding the optimal model for predicting the actual radioactivity. This means that the desired characteristics of candidate models, such as type of data and linearity, were identified in advance, and we had specific suitable models in mind. The correlation and linear relationship in the linear

regression statistical analysis results support this. For modeling, the RF and XGBoost algorithms are already widely known to be suitable for regression and classification [15–17]. RF, which performed best in this study, has many advantages: it is fast in both model training and evaluation, is robust to outliers, can capture complex nonlinear associations, and has been shown to handle challenges arising from small data sets [18–20]. RF performed very well even with a small data set in this study. We need to focus on how much the concern about radiation exposure will be improved. Several studies have been conducted on reducing the exposure of workers to radiation safely. Among these, L tzen U et al. developed a shielded measurement method by creating calibration curves using measurements of shielded I-131 capsules from a dose calibrator and a well-type and a thyroid uptake probe. This method reduced the effective radiation dose by 94.9% [21]. We calculated the equivalent dose to the hand when directly measuring the radioactivity of 7400 MBq of I-131 in the introduction of this study. If our trained RF model were used, although there would be the disadvantage of increased working time due to the measurement of the external dose rates, an equivalent dose of about 0.00885 mSv would be expected, which would represent a 94% reduction compared to the direct measurement method (Table 5). This usage of a predictive model to reduce the equivalent dose is not limited to determining the radioactivity of high-dose I-131, but can be extend to predicting the radioactivity

**Table 4.** Accuracy the models

Model	RMSE	MSE	MAE	$R^2$
Linear regression	268.145	71901.870	231.678	0.917
DT	108.890	11856.920	19.237	0.986
RF	8.894	79.098	6.546	0.999
XGBoost	10.209	104.219	7.676	0.999

**Table 5.** Comparison of equivalent doses

	Working time (sec)	Distance to hand (m)	Gamma constant (mSv*MBq/m <sup>2</sup> /h)	Equivalent dose (mSv)	Total (mSv)
Direct measurement	10	0.1	7.638E-05	0.157000	0.157
Predictive model using dose rate measurement	60	0.1	5.784E-07	0.007134	0.00885
	60	0.3	1.108E-06	0.001518	
	60	1.0	1.603E-06	0.000198	

of several sources using measured external dose rates, and further research is expected. The limitations of this study are as follows. First, overfitting could not be excluded to some extent in the two ensemble models. In the training set, the RMSE was 8,894 for the RF model and 10,209 for the XGBoost model, but when the performance of the RF and XGBoost models was validated using 11 newly measured sets, the RMSE was 38,726 for the RF model and 42,670 for the XGBoost model (Table 6). According to Breiman L, a relatively large number of variables are required to get a near-optimal test set error [22]. This means that external dose rate values measured at various distances are required, and increasing the number of these variables would improve the performance of the model. The performance of the model could be improved if more measurement distances were added in place of the measurement at 1 m, where the variable importance is 0 (Table 7), but this has a drawback in that the measurement time increases. These issues should be considered going forward. For

therapeutic administrations of I-131, if more than a 10% variance from the prescribed dose or dosage range occurs, the referring physician and patient should be notified [23, 24]. The second limitation is that the model cannot predict when this variance of 10% or more will occur. When RF is employed for regression, it is unable to predict doses beyond the range seen in the training data set. Because, in the training data set, there were no cases that had an error of more than 10%, the model was not able to learn to predict such cases. We expect this problem can be addressed by employing a classification model such as support vector machine (SVM), which is capable of learning to identify cases where there is a 10% or greater variance in dose.

## V. Conclusion

To predict the radioactivity of high-dose I-131 using a supervised machine learning model, we collected a

**Table 6.** Example of validating trained random forest and extreme gradient boost models

	Dose rate ( $\mu\text{Sv/h}$ )			Actual activity (MBq)	Predictive model activity (MBq)	
	1.0 m	0.3 m	0.1 m		RF	XGBoost
	5.83	45.40	216.00	3940.50	3925.70	3922.00
	5.80	48.10	226.00	3907.20	3947.90	3947.90
	4.50	38.50	175.00	3833.20	3822.10	3807.30
	5.20	46.80	186.00	3899.80	3907.20	3881.30
	5.90	43.70	193.00	3903.50	3888.70	3873.90
	5.54	41.10	217.00	3899.80	3888.70	3914.60
	5.61	44.70	216.00	3907.20	3918.30	3914.60
	5.52	43.50	213.00	3896.10	3903.50	3903.50
	5.98	43.10	197.00	3966.40	3888.70	3873.90
	6.29	42.80	191.00	3925.70	3925.70	3929.40
	5.53	44.00	211.00	3822.10	3910.90	3907.20
			RMSE		38,726	42,670
			MSE		1499,677	1820,770
			MAE		25,900	31,282

**Table 7.** Random forest variable importance

Variable	Overall
0.3 m	100.000
0.1 m	16.710
1.0 m	0.000

data set containing measurements of external dose rates from a shielded I-131 at various distances. We found that there was a correlation and a linear relationship between the external dose rate of high-dose I-131 and the actual radioactivity. Random forest achieved the best predictive capability with an RMSE of 8,894, MSE of 79,098, and MAE of 6,546. Improving the model's performance in the future is expected to contribute to lowering exposure among radiology technologists.

## REFERENCES

- [1] Kopisch A, Martin CB, Grantham V. Exposure to technologists from preparing and administering therapeutic 131I: How frequently should we bio-assay? *J Nucl Med Technol*. 2011;39:60-2.
- [2] Jang DG, Yang SO, Kim JG, Lee SH, Choi HS, Bae CW. A study on the shielding of iodine 131 using monte carlo simulation. *Journal of Radiological Science and Technology*. 2014;37(2):143-50.
- [3] Geoffrey MC. Intelligent Imaging: Artificial Intelligence Augmented Nuclear Medicine. *J Nucl Med Technol*. 2019;47:217-22.
- [4] Tang A, Tam R, Cadrin-Ch nevert A, Guest W, Chong J, Barfett J, et al. Canadian association of radiologists white paper on artificial intelligence in radiology. *Can Assoc Radial J*. 2018;69:120-35.
- [5] Lundervold AS, Lundervold A. An overview of deep learning in medical imaging focusing on MRI. *Z Med Phys*. 2019;29:102-27.
- [6] Tukey JW. *Exploratory data analysis*. Pearson, 1977;2:131-60.
- [7] Berrar D. Cross-validation. *Encyclopedia of Bioinformatics and Computational Biology*. 2018;1:542-5.
- [8] Nam YW, Arai Y, Kunizane T, Koizumi A. Water leak detection based on convolutional neural network using actual leak sounds and the hold-out method. *Water Supply*. 2021;21(7):3477-85.
- [9] Lynch CM, Abdollahi B, Fuqua JD, De Carlo AR, Bartholomai JA, Balgemann RN, et al. Prediction of lung cancer patient survival via supervised machine learning classification techniques. *International Journal of Medical Informatics*. 2017;108:1-8.
- [10] Muhammad LJ, Salisu S, Yakubu A, Malgwi YM, Abdullahi ET, Mohammed IA, et al. Using decision tree data mining algorithm to predict causes of road traffic accidents, its prone locations and time along Kano-Wudil highway. *Int J Database Theory Appl*. 2017;10(11):197-208.
- [11] Muhammad LJ, Algehyne EA, Usman SS, Ahmad A, Chakraborty C, Mohammed IA. Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. *SN Computer Science*. 2021;2(1):11.
- [12] Muhammad LJ, Ali AG, Zakariyau YB, Mohammed IA. Multi query optimization algorithm using semantic and heuristic approaches. *Int J Database Theory Appl*. 2016;6(9):219-26.
- [13] Hastie T, Tibshirani R, Friedman J. *Random forests*. In: *The Elements of Statistical learning: Data mining, inference, and prediction*, 2nd ed. Springer Series in Statistics. New York: Springer; 2009:587-604.
- [14] Khandelwal N. A brief introduction to XGBoost [Internet]. *Towards Data Science*; 2020 [cited 2020 Jul 7]. Available from: <https://towardsdatascience.com/a-brief-introduction-to-XGBoost-3eae2e3e5d6>
- [15] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. *ACM*. 2016:785-94.
- [16] Belgiu M, Dragut L. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2016;114:24-31.
- [17] Speiser JL, Miller ME, Tooze EJ. A comparison of random forest variable selection methods for classification prediction modeling. *Expert Syst Appl*. 2019;134:93-101.
- [18] Han S, Kim H, Lee YS. Double random forest. *Mach Learn*. 2020:1569-86.
- [19] Qi Y. Random forest for bioinformatics. *Computer Science*. 2012:307-23.
- [20] Han S, Williamson BD, Fong Y. Improving random forest predictions in small datasets from two-phase sampling designs. *BMC Med Inform Decis Mak*.



- 2021;322(21).
- [21] L tzen U, Zhao Y, Marx M, Imme T, Assam I, Siebert FA, et al. Effective method of measuring the radioactivity of [131I]-capsule prior to radio-iodine therapy with significant reduction of the radiation exposure to the medical staff. *J Appl Clin Med Phys*. 2016;17(4):59-72.
- [22] Breiman L. Random forests. *Machine Learning*. 2001;45:5-32.
- [23] Levine G, Malhi B, Struttman L. Discrepancies in dose calibrator assays for various forms of therapeutic Iodine-131. *J Nucl Med Technol*. 1984;12(2):84-6.
- [24] Kang SJ, Lee DH, Soh Y, Lee JW. Evaluation of caregivers' exposed dose and patients' external dose rate for radioactive iodine (I-131) therapy administration in isolated ward. *Journal of Radiological Science and Technology*. 2022;45(4):347-53.

구분	성명	소속	직위
제1저자	유연욱	국립암센터 핵의학과	방사선사
공동저자	이충운	국립암센터 핵의학과	방사선사
교신저자	김정수	동남보건대학교 방사선학과	부교수