MATHEMATICAL MODELING AND COMPUTING, Vol. 11, No. 1, pp. 230–238 (2024)

Machine learning models selection under uncertainty: application in cancer prediction

Lamrani Alaoui Y.[1], Benmir M.[2], Aboulaich R.[2]

[1] *Mohammadia School of Engineering (EMI), Mohammed V University in Rabat, Rabat, Morocco*
[2] *Mohammadia School of Engineering, Mohammed V University in Rabat, Rabat, Morocco*

(Received 7 July 2023; Revised 3 March 2024; Accepted 4 March 2024)

Cancer stands as the foremost global cause of mortality, with millions of new cases diagnosed each year. Many research papers have discussed the potential benefits of Machine Learning (ML) in cancer prediction, including improved early detection and personalized treatment options. The literature also highlights the challenges facing the field, such as the need for large and diverse datasets as well as interpretable models with high performance. The aim of this paper is to suggest a new approach in order to select and assess the generalization performance of ML models in cancer prediction, particularly for datasets with limited size. The estimates of the generalization performance are generally influenced by numerous factors throughout the process of training and testing. These factors include the impact of the training–testing ratio as well as the random selection of datasets for training and testing purposes.

Keywords: *cancer prediction; machine learning; hesitant fuzzy logic; MCDM.*
2010 MSC: 03B52, 68T37, 62-07 **DOI:** 10.23939/mmc2024.01.230

1. Introduction

Machine learning has emerged as a promising tool for cancer prediction and diagnosis in recent years, with a plethora of studies showcasing its potential in various cancer types [1–5]. Traditional methods of cancer diagnosis, such as biopsy and radiography, can be invasive, time-consuming, and sometimes inconclusive. However, machine learning techniques can analyze complex data patterns and identify subtle changes in the human body, leading to earlier and more accurate cancer detection.

However, despite the promising results of ML, there are still challenges in applying machine learning to cancer prediction and diagnosis. One of the major challenges is the lack of large-scale, high-quality datasets, which are essential for training accurate machine learning models. Additionally, the interpretability of machine learning models in the context of cancer diagnosis is still an open research question.

When it comes to cancer prediction in small datasets, assessing the performance of a machine learning model becomes challenging. In order to overcome the limitations of standard approaches and improve the evaluation process of Machine Learning models, this study aims to suggest a novel framework for model selection in cancer prediction based on a Hesitant Fuzzy MCDM methods.

The remaining sections of the paper are organized as follows: in Section 2, we provide a concise explanation of the significance of MCDM in performance evaluation of ML models; in Section 3, we present our proposed approach; in Section 4, we introduce the Hesitant fuzzy TOPSIS and Hesitant fuzzy VIKOR; in Section 5, we demonstrate the application of our suggested approach to select an optimal ML model for cancer prediction. Finally, in Section 6, we conclude the study.

2. Performance assessment of machine learning models

The assessment and selection of ML algorithms is a very interesting research topic in data science [6,7]. One way to assess the prediction quality of a machine learning model is through a training testing strategy. The initial data is to be split into training and test set. The training set will be used for

hyper-parameters tuning and models' training and then the test set is to be used for the model's evaluation.

The literature highlights a panoply of performance measures for ML classifiers [8]. These criteria include: accuracy score, precision, recall, F1-score and area under curve (AUC).

Table 1. Confusion matrix.

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

These criteria are based mainly on the confusion matrix (Table 1) and are commonly used by researchers and practitioners. Let TP, FP, TN and FN stand for:

- True positive (TP) refers to the instances where a model correctly predicts a positive outcome;
- False positive (FP)occurs when a model incorrectly predicts a negative outcome as positive;
- True negative (TN) refers to the instances where a model correctly predicts a negative outcome;
- False negative (FN)occurs when a model incorrectly predicts a positive outcome as negative.

The aforementioned criteria are defined as follows:

— **Accuracy**: It is the number of observations correctly identified to the total number of observations. This ratio is recommended when we have balanced classes; which means there is an equal number of observations belonging to each class.

— **Precision**: It represents the percentage of predicted positive values that are really positive. The precision measure is recommended when the cost associated with the false positive is high. For example, if a sick patient is predicted as not sick, the risk will be very high if the sickness is contagious.

— **Sensitivity**: Also known as recall, or the true positive rate (TPR). It represents the percentage of actual positive values that are predicted positive. This measure must be chosen when the cost associated with the false negative is high.

— **F1 score**: This metrics is used when we seek a balance between both the recall and precision, its formula is as follows:

$$2 \times \frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}}$$

— **Area under curve**: The area under curve (AUC): the AUC range is from 0 to 1. The higher the value of AUC is, the higher the performance of the model; the AUC is given by

$$\text{AUC} = \frac{1 + \text{TPR} - \text{FPR}}{2}.$$

In practice, generally, there is no classifier that has the best results across all performance metrics. So, it is slightly difficult to select the best ML model. According to Kou [9], the selection of machine learning algorithms can be seen as a MCDM problem that involves more than one criterion.

MCDM techniques are a mathematical approach designed to help make decisions in the presence of multiple and conflicting criteria. The idea of MCDM is to define a set of objectives or alternatives, select criteria to assess these objectives, assign weights for criteria, and then apply an algorithm to rank and classify alternatives [10].

Let $M_1, M_2, \ldots, M_n$ be a set of alternatives to be classified with respect to a set of criteria $C_1, C_2, \ldots, C_m$, and $W = (w_1, w_2, \ldots, w_m)$ be the weighs' vector of all criteria. A MCDM problem is defined as

$$\begin{pmatrix} r_{11} & r_{12} & \ldots & r_{1m} \\ r_{21} & r_{22} & \ldots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \ldots & r_{nm} \end{pmatrix},$$

where r_{ij} is the rating of the i^{th} alternative with respect to the j^{th} criterion. To solve such a kind of problem, a panoply of Multi-Criteria-Decision-Making (MCDM) methods were developed such as

TOPSIS and VIKOR. These methods were first developed to hand quantitative data, then they were extended to cover qualitative, subjective, or uncertain data as well.

In classical MCDM techniques, the alternatives' ratings r_{ij} are supposed known precisely, however in many real-world situations, including the evaluation of a ML model, the ratings may be imprecise and uncertain. Thus, fuzzy MCDM methods may be a practical solution.

In many real-world applications, which include cancer prediction, obtaining a large dataset poses a significant challenge. There is a strong interest in exploring effective strategies to maximize the utility of datasets with limited size [7]. Numerous approaches have been suggested to evaluate the performance of machine learning (ML) models, with a primary focus on the size of the available dataset. The holdout method [11] is among the widely adopted approaches. Despite its simplicity in programming and speedy execution, the holdout method is statistically less robust. Whenever the initial dataset is randomly divided into two parts, there is a possibility of altering the sample statistics [12] and then we may obtain different results of the performance estimates [13]. The generalization performance of a machine learning model is additionally influenced by the training-testing ratio [6, 11].

The objective of this study is to propose an innovative approach for evaluating the predictive performance of machine learning (ML) algorithms while considering the uncertainty inherent in performance estimates. Our primary focus is on addressing the uncertainty stemming from both the training-testing ratio and the random selection of training and testing data.

3. The proposed approach

In order to provide a comprehensive assessment of a machine learning model, we propose assessing its performance using three distinct training-testing ratios (Table 2). We repeat the train-test splitting procedure ten times for each ratio and subsequently calculate the average results of each performance estimate.

Table 2. Commonly used training-testing ratios.

Training dataset	Testing dataset
60%	40%
70%	30%
80%	20%

A model M_i performance according to a performance estimate C_j is then given by three possible values $p_{ij} = (\gamma_{ij}^1, \gamma_{ij}^2, \gamma_{ij}^3)$ each value represents the obtained score according to a specific training-testing ratio. To calculate the different values of p_{ij} the following steps are adopted:

1. Gain insights into the data and perform necessary cleaning processes;
2. Select relevant variables for the model;
3. If required, address data imbalance issues that may negatively affect model training;
4. Repeat the following steps k times:
   — Separate the data randomly into training and testing sets using a predefined training-testing ratio, denoted as $\alpha \in]0, 1[$;
   — Optimize the hyperparameters of the model using Cross validation on the training data.
   — After identifying the optimal hyperparameters, train the final optimized classifier using the complete training dataset.
   — Assess the predictive performance of the machine learning model on the testing data.
5. Calculate the performance γ_{ij} of the model M_i based on a specific criterion C_j and a predetermined training-testing ratio by taking the average of the scores acquired from the k repetitions.
6. Repeat steps 4 and 5 for three distinct values of α. The model M_i performance according to C_j is now given by $p_{ij} = (\gamma_{ij}^1, \gamma_{ij}^2, \gamma_{ij}^3)$. 3 is the number of distinct training-testing ratios.
7. Construct the decision matrix using hesitant fuzzy numbers.

The Multiple Criteria Decision Making (MCDM) problem can be represented as shown in Table 3. Where $M_1, M_2, \ldots, M_n$ is a set of machine learning models to be classified with respect to a set of criteria $C_1, C_2, \ldots, C_m$.

The performance $p_{ij} = (\gamma_{ij}^1, \gamma_{ij}^2, \gamma_{ij}^3)$ of a model M_i with respect to a specific criterion C_j can be seen as an hesitant fuzzy number. A hesitant fuzzy (HFN) [14] offers a more comprehensive representation of fuzzy and uncertain information, it is expressed as

$$h = (\gamma_1, \gamma_2, \gamma_3),$$

where $(\gamma_1, \gamma_2, \gamma_3)$ is a set of ratings in $[0, 1]$ denoting the possible membership degrees of an element $x \in X$ to a given set A. Under hesitant fuzzy information, the decision matrix is given in Table 4, where $h_{ij} = (\gamma_{ij}^1, \gamma_{ij}^2, \gamma_{ij}^3)$ is a Hesitant Fuzzy Number (HFN) that represents the evaluation of the algorithm M_i with respect to the performance estimate C_j according to 3 different training testing ratios.

Table 3. The proposed decision matrix.

Model	C_1	...	C_m
M_1	p_{11}	...	p_{1m}
M_2	p_{21}	...	p_{2m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
M_n	p_{n1}	...	p_{nm}

To select the best ML model, we need to solve the hesitant fuzzy decision-making problem using a MCDM method like hesitant fuzzy TOPSIS or hesitant fuzzy VIKOR [15]. VIKOR operates by selecting a compromise solution from a set of alternatives, aiming to maximize group utility and minimize individual regret. On the other hand, TOPSIS identifies a solution that has the shortest distance to the ideal solution and the farthest distance to the negative-ideal solution. Both TOPSIS and VIKOR can be employed to rank the predictive performance of a collection of ML models and subsequently determine the optimal choice.

Table 4. Decision matrix under hesitant fuzzy information.

Model	C_1	...	C_m
M_1	h_{11}	...	h_{1m}
M_2	h_{21}	...	h_{2m}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
M_n	h_{n1}	...	h_{nm}

Hesitant fuzzy logic is an expansion of classical fuzzy logic [16] used to address situations where a single element can have a range of possible membership values [14]. It considers all possible values instead of using an aggregation operator to obtain a single value.

Definition 1. *A hesitant fuzzy set B is defined as $B = \{\langle x, h_B(x)\rangle | x \in X\}$, where $h = h_B(x)$ is a HFN comprising a set of membership degrees in $[0, 1]$ indicating the potential degrees of membership of element $x \in X$ to B. A HFN can be represented as $h = (\gamma_1, \gamma_2, \ldots, \gamma_k)$.*

Consider three HFNs denoted as h, h_a, h_b, and $\lambda > 0$, some arithmetic operations are defined as follows:

1. The complement of h is given as $h^c = \cup_{\gamma \in h}\{1 - \gamma\}$;
2. $h^\lambda = \cup_{\gamma \in h}\{\gamma^\lambda\}$;
3. $\lambda h = \cup_{\gamma \in h}\{1 - (1 - \gamma)^\lambda\}$;
4. $h_a \cup h_b = \cup_{\gamma_1 \in h_a, \gamma_2 \in h_b} \max\{\gamma_1, \gamma_2\}$;
5. $h_a \cap h_b = \cup_{\gamma_1 \in h_a, \gamma_2 \in h_b} \min\{\gamma_1, \gamma_2\}$;
6. $h_a \oplus h_b = \cup_{\gamma_1 \in h_a, \gamma_2 \in h_b} \{\gamma_1 + \gamma_2 - \gamma_1\gamma_2\}$;
7. $h_a \otimes h_b = \cup_{\gamma_1 \in h_a, \gamma_2 \in h_b} \{\gamma_1\gamma_2\}$;
8. The first hesitant Hamming distance between h_a and h_b is defined as [15]:

$$D_1(h_a, h_b) = \|h_a - h_b\| = \frac{1}{k}\sum_{l=1}^{k}\left|h_a^{\sigma(l)} - h_b^{\sigma(l)}\right|$$

   where $h_a^{\sigma(l)}(l = 1, 2, \ldots, k)$ and $h_b^{\sigma(l)}$ $(l = 1, 2, \ldots, k)$ are the l^{th} smallest value in h_a and h_b respectively.
9. The second Hamming distance between h_a and h_b under hesitant information is defined as [17]:

$$D_2(h_a, h_b) = \frac{1}{2}(g(h_a, h_b) + g(h_a, h_b)),$$

   where $g(h_i, h_j) = \frac{1}{k_{h_i}}\sum_{\gamma_i \in h_i} \min_{\gamma_j \in h_j} \|\gamma_i - \gamma_j\|$.

4. Hesitant fuzzy MCDM approaches

In order to enhance decision-making processes, numerous models and tools have been developed. An illustrious example in this regard is the Multi-Criteria Decision Making (MCDM) framework. It involves employing a scientific approach to make decisions that take multiple criteria into account [18]. The fundamental principle of MCDM lies in dissecting a problem into smaller components (alternatives, criteria, etc.) and subsequently establishing a hierarchical order among the choices, thereby enabling a comprehensive mathematical understanding of the problem. Within the scope of this study, we undertake a comparative analysis of two methodologies, namely TOPPSIS and VIKOR, both applied in the context of hesitant fuzzy information. These methods are highly adaptable, easy to comprehend, and possess a robust mathematical foundation when compared to various other MCDM approaches [19,20].

Let $M_1, M_2, \ldots, M_n$ be a set of alternatives to be evaluated with respect to a set of criteria $C_1, C_2, \ldots, C_n$ and let $W = (w_1, w_2, \ldots, w_m)$ be the weight vector of all criteria. A hesitant fuzzy decision matrix is given by

$$\begin{pmatrix} h_{11} & h_{12} & \ldots & h_{1m} \\ h_{21} & h_{22} & \ldots & h_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \ldots & h_{nm} \end{pmatrix},$$

where h_{ij} is a Hesitant Fuzzy Number (HFN) which represents the score of the alternative A_i with regards to the criterion C_j as perceived by the decision makers.

According to [21], the TOPSIS method, when applied in the context of hesitant fuzzy information, can be succinctly outlined through the following steps:

1. Determine the hesitant positive ideal solution (HPIS) I^+ and the hesitant negative ideal solution (HNIS) I^-:

$$I^+ = \{h_1^*, h_2^*, \ldots, h_n^*\},$$

where

$$h_j^* = \max\left\{h_{ij}^{\sigma(k)} \middle| i = 1, 2, \ldots, m\right\} = \left\{(h_j^1)^+, (h_j^2)^+, \ldots, (h_j^l)^+\right\}, \quad I^- = \left\{h_1^-,\ h_2^-, \ldots, h_n^-\right\},$$

$$h_j^- = \min\left\{h_{ij}^{\sigma(k)} \middle| i = 1, 2, \ldots, m\right\} = \left\{(h_j^1)^-, (h_j^2)^-, \ldots, (h_j^l)^-\right\},$$

$h_{ij}^{\sigma(k)}$ is the k^{th} smallest value in h_{ij}.

2. Calculate the separation measure ${D_i}^+$ and ${D_i}^-$ of each alternative from the HPIS and HNIS, respectively using hesitant Hamming distance:

$$D_i^+ = \sum_{j=1}^{n} w_j \|h_{ij} - h_j^*\|, \quad i = 1, 2, \ldots, m,$$

$$D_i^- = \sum_{j=1}^{n} w_j \|h_{ij} - h_j^-\|, \quad i = 1, 2, \ldots, m,$$

where w_j denotes the weight of j^{th} criterion.

3. To compute the relative closeness coefficient C_i for each alternative with respect to the hesitant ideal solution, use the following formula:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-}.$$

The alternatives are then ranked according to C_i values. The best alternative is that with the highest C_i value.

When it comes to hesitant fuzzy VIKOR, its development process can be summarized as follows [15]:

1. Identify the HPIS I^+ and the hesitant HNIS I^-.

2. Compute the values of R_i and S_i

$$S_i = \sum_{j=1}^{n} w_j \frac{\|h_{ij} - h_j^*\|}{\|h_j^- - h_j^*\|}, \quad i = 1, 2, \ldots, m;$$

$$R_i = \max_j w_j \frac{\|h_{ij} - h_j^*\|}{\|h_j^- - h_j^*\|}, \quad i = 1, 2, \ldots, m.$$

3. Calculate the Q_i values

$$Q_i = \nu \frac{(S_i - S^+)}{(S^- - S^+)} + (1 - \nu) \frac{(R_i - R^+)}{(R^- - R^+)}, \quad i \in \{1, 2, \ldots, m\},$$

where

$$S^+ = \max_i S_i, \quad S^- = \min_i S_i,$$

$$R^+ = \max_i R_i, \quad R^- = \min_i R_i,$$

and ν represents the maximum group utility, typically set to 0.5 [22].

4. Arrange the alternatives in descending order based on the values of S, R, and Q. This process will yield three separate ranking lists.
5. If the following two conditions are met, the compromise solution M' is the alternative with the minimum value according to the Q measure:
   — C_1: Acceptable advantage: $Q(M'') - Q(M') \geqslant \frac{1}{m-1}$, M'' is the second-best alternative in terms of Q_i and m is the number of alternatives being compared.
   — C_2 : Acceptable stability in decision making: The alternative M' must also be the best in terms of S and R.
6. If one of two conditions is not satisfied, a set of compromise solutions is proposed, which includes:
   — Select the alternatives M' and M'' if only condition C_2 is not satisfied.
   — Select the alternatives $M', M'', \ldots M^{(k)}$ if the condition C_1 is not satisfied. The alternative $M^{(k)}$ is determined based on the relation $Q(M^{(k)}) - Q(M') < \frac{1}{m-1}$.

5. Application in cancer prediction

In this section, the proposed approach is applied in order to compare the performance of different machine learning models in cancer prediction. The ML models are trained with the Breast cancer Wisconsin (diagnostic) dataset from scikit-learn library (Table 5).

Table 5. Description of Breast cancer Wisconsin dataset.

Number of Instances	569
Number of Attributes of variables	30 numeric, predictive variables and the class
Missing variable Values	None
Class Distribution	212 – Malignant 357 – Benign

It is a widely used benchmark dataset for breast cancer classification. The features are computed from digitized images of breast mass samples and represent quantitative measures of their morphological characteristics. The dataset is labeled with binary outcomes, where 1 indicates malignant (cancerous) and 0 indicates benign (non-cancerous) cases. The goal of using this dataset is to train machine learning models to accurately classify breast masses as benign or malignant based on the provided features. This dataset is widely used in research and educational settings to study and develop breast cancer classification algorithms and assess their performance.

Five ML algorithms are trained using python scikit-learn library and according to the suggested approach. The performances of the different models are assessed based on four criteria: accuracy, precision, recall and F1-score. These algorithms are Logistic regression, Decision tree classifier, support vector machine (SVM), K-Nearest Neighbors (KNN) and Random Forest.

Logistic regression (LR) is a generalized linear regression classifier very used in binary classification. LR uses the logistic, called also sigmoid function to predict the probability of belonging to a default class. The assumptions of logistic regression are quite similar to those of linear regression.

The Support vector machine (SVM) is a classifier used to model linear and non-linear phenomena. The fundamental idea is to look for the best hyperplane to separate two or more classes based on training data.

The K-Nearest Neighbors (KNN) is one of the simplest and well-known classifiers. Its main idea is to use all training data to predict the outcome of unlabeled data based on a similarity measure (e.g., Euclidian distance). Each new point is assigned to the most frequent category among its K nearest neighbors.

Decision tree classifier is a fast classification and regression algorithm that provides a simple visualization and interpretation of data patterns. However, over-fitting is a significant feature of this algorithm. One of several solutions to overcome this issue is the ensemble learning. The basic idea is to develop a model that makes predictions based on a combination of multiple individual models. In this study Random Forest algorithm was implemented.

Random forest is a popular ensemble-learning algorithm that integrates a large number of decision tree classifiers and then selects the optimal solution by means of voting.

The obtained results are summarized in Table 6. One can notice that there is no big difference between the performances of the used machine learning classifiers. That may be explained by the high-quality of the data, its balanced classes, and its highly relevant features for classifying breast cancer as malignant or benign.

Table 6. Hesitant fuzzy decision matrix for the adopted machine learning classifiers.

Model	Accuracy	Precision	Recall	F1 Score
SVM	(.95, .95, .96)	(.94, .95, .96)	(.97, .97, .98)	(.96, .96, .97)
Random forest	(.95, .96, .97)	(.94, .96, .97)	(.97, .97, .98)	(.96, .96, .97)
Logistic regression	(.94, .94, .95)	(.93, .94, .95)	(.97, .97, .98)	(.95, .95, .96)
CART	(.92, .92, .93)	(.92, .92, .95)	(.94, .95, .96)	(.94, .94, .95)
KNN	(.93, .93, .94)	(.92, .92, .93)	(.97, .97, .97)	(.95, .95.95)

To select the best machine learning model, extended versions of both TOPSIS and VIKOR methods are implemented [15]. The weights of criteria are also calculated using an extended version of the entropy measure under hesitant environment [17]. Let h, a hesitant fuzzy number, the hesitant fuzzy entropy measure is given by

$$E(h) = S(h, h^c) = 1 - D_2(h, h^c),$$

where $h^c = \cup_{\gamma \in h}\{1 - \gamma\}$. The criteria weight is obtained as follows:

$$w_j = \frac{1 - E_j}{m - \sum_j^m E_j}, \quad j = 1, \ldots, m,$$

where $E_j = \frac{1}{n}\sum_{i=1}^{n} E(h_{ij})$, $j = 1, \ldots, m$.

The obtained criteria weights as well as the results of both TOPSIS and VIKOR are given in Tables 7, 8 and 9 respectively.

Table 7. Obtained weights of criteria.

Accuracy	Precision	Recall	F1 Score
0.245	0.245	0.259	0.251

Table 7 shows that the weights affected to the different criteria are quite similar. Tables 8 and 9 suggest that the random forest classifier is the best algorithm, followed by the SVM, and then KNN.

Table 8. Results of the Hesitant fuzzy VIKOR.

Model	S rank	R rank	Q rank	Final rank
Random forest	0.017	0.01	0	1
SVM	0.182	0.112	0.372	—
Logistic regression	0.2	0.172	0.516	—
KNN	0.273	0.234	0.717	—
CART	0.531	0.25	1	—

Table 9. Results of the Hesitant fuzzy TOPSIS.

Model	D^+	D^-	C	Final rank
Random forest	0.017	0.580	0.971	1
SVM	0.004	0.014	0.791	2
Logistic regression	0.571	0.571	0.500	4
KNN	0.008	0.012	0.587	3
CART	0.563	0.013	0.023	5

6. Conclusion

The aim of this paper was to introduce a new framework for selecting machine learning models in cancer prediction under fuzzy environment. The adopted approach could be more practical as it considers the uncertainty associated with performance estimates, arising from the training-testing ratio and the random selection of datasets. By acknowledging and addressing these uncertainties head-on, this framework opens the door to a more robust methodology for assessing ML models in cancer prediction.

[1] Zhang C., Hu J., Li H., Ma H., Othmane B., Ren W., Yi Z., Qiu D., Ou Z., Chen J., Zu X. Emerging biomarkers for predicting bladder cancer lymph node metastasis. Frontiers in Oncology. **11**, 648968 (2021).

[2] Wang P., Li Y., Reddy C. K. Machine learning for survival analysis: A survey. ACM Computing Surveys. **51** (6), 1–36 (2019).

[3] Levine A. B., Schlosser C., Grewal J., Coope R., Jones S. J. M., Yip S. Rise of the machines: advances in deep learning for cancer diagnosis. Trends in Cancer. **5** (3), 157–169 (2019).

[4] Huang S., Yang J., Fong S., Zhao Q. Artificial intelligence in cancer diagnosis and prognosis: Opportunities and challenges. Cancer letters. **471**, 61–71 (2020).

[5] Abreu P. H., Santos M. S., Abreu M. H., Andrade B., Silva D. C. Predicting breast cancer recurrence using machine learning techniques: a systematic review. ACM Computing Surveys. **49** (3), 1–40 (2016).

[6] Nguyen Q. H., Ly H.-B., Ho L. S., Al-Ansari N., Le H. V., Tran V. Q., Prakash I., Pham B. T. Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. Mathematical Problems in Engineering. **2021**, 4832864 (2021).

[7] Witten I. H., Frank E., Hall M. A. Credibility: evaluating what's been learned. Data Mining: Practical Machine Learning Tools and Techniques. 147–187 (2011).

[8] Japkowicz N., Shah M. Performance evaluation in machine learning. Machine Learning in Radiation Oncology. 41–56 (2015).

[9] Kou G., Lu Y., Peng Y., Shi Y. Evaluation of classification algorithms using MCDM and rank correlation. International Journal of Information Technology & Decision Making. **11** (01), 197–225 (2012).

[10] Qu Z., Wan C., Yang Z., Lee P. T.-W. A discourse of multi-criteria decision making (MCDM) approaches. Multi-Criteria Decision Making in Maritime Studies and Logistics. 7–29 (2018).

[11] Uçar M. K., Nour M., Sindi H., Polat K. The effect of training and testing process on machine learning in biomedical datasets. Mathematical Problems in Engineering. **2020**, 2836236 (2020).

[12] Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. Preprint arXiv:1811.12808 (2018).

[13] Zheng A. Evaluating machine learning models: a beginner's guide to key concepts and pitfalls. O'Reilly Media (2015).

[14] Torra V. Hesitant fuzzy sets. International Journal of Intelligent Systems. **25** (6), 529–539 (2010).

[15] Zhang N., Wei G. Extension of VIKOR method for decision making problem based on hesitant fuzzy set. Applied Mathematical Modelling. **37** (7), 4938–4947 (2013).

[16] Zadeh L. A. Fuzzy sets. Information and Control. **8** (3), 338–353 (1965).

[17] Hu J., Zhang X., Chen X., Liu Y. Hesitant fuzzy information measures and their applications in multi-criteria decision making. International Journal of Systems Science. **47** (1), 62–76 (2016).

[18] Gal T., Stewart T., Hanne T. (Eds.). Multicriteria decision making: advances in MCDM models, algorithms, theory, and applications. Springer Science + Business Media, New York (2013).

[19] Hwang C. L., Yoon K. Methods for multiple attribute decision making. Multiple Attribute Decision Making. 58–191 (1981).

[20] Shih H.-S., Shyur H.-J., Lee E. S. An extension of TOPSIS for group decision making. Mathematical and Computer Modelling. **45** (7–8), 801–813 (2007).

[21] Xu Z., Zhang X. Hesitant fuzzy multi-attribute decision making based on TOPSIS with incomplete weight information. Knowledge-Based Systems. **52**, 53–64 (2013).

[22] Sayadi M. K., Heydari M., Shahanaghi K. Extension of VIKOR method for decision making problem with interval numbers. Applied Mathematical Modelling. **33** (5), 2257–2262 (2009).

Вибір моделей машинного навчання в умовах невизначеності: застосування в прогнозуванні раку

Ламрані Алауї Ю.[1], Бенмір М.[2], Абулайх Р.[2]

[1] *Інженерна школа Мохаммадіа (EMI), Університет Мухаммеда V у Рабаті, Рабат, Марокко*
[2] *Інженерна школа Мохаммадіа, Університет Мухаммеда V у Рабаті, Рабат, Марокко*

Рак є головною глобальною причиною смертності, з мільйонами нових випадків діагностування щороку. Багато дослідницьких статей обговорювали потенційні переваги машинного навчання (ML) у прогнозуванні раку, включаючи покращене раннє виявлення та персоналізовані варіанти лікування. У літературі також висвітлюються проблеми, що постають перед цією сферою, наприклад, потреба у великих і різноманітних наборах даних, а також високоефективних інтерпретованих моделях. Метою цієї статті є запропонувати новий підхід до вибору та оцінки ефективності узагальнення моделей ML у прогнозуванні раку, особливо для наборів даних обмеженого розміру. На оцінки ефективності узагальнення, як правило, впливають численні фактори протягом усього процесу навчання та тестування. Ці фактори включають вплив співвідношення навчання та тестування, а також випадковий вибір наборів даних для цілей навчання та тестування.

Ключові слова: *прогноз раку; машинне навчання; нечітка логіка; MCDM.*