https://doi.org/10.31891/csit-2024-1-1
UDC 004

ABUBAKAR Sadiq Abdulhameed
Ternopil Ivan Puluj National Technical University, Ternopil Ukraine
SVERSTIUK Andriy
Ternopil Ivan Puluj National Technical University, Ternopil Ukraine,
Ternopil Ivan Horbachevsky National Medical University, Ternopil, Ukraine

ANALYTICAL REVIEW OF PUBLICATIONS ON MACHINE LEARNING METHODS IN ONCOLOGY AND APPROACH TO EVALUATING THEIR QUALITY

The work includes an analytical review of publications on machine learning methods in oncology and an approach to evaluating their quality. An analysis of publications by year was conducted in the Web of Science and Scopus bibliometric databases. The highest number of authors, the number of publications among universities, the number of countries, and publication categories in the Scopus bibliometric database on machine learning methods in oncology are presented. A multifactor regression prediction model for bone tissue density in oncological pathology predicting four severity grades of the studied disease course was proposed. This model included the following factors with corresponding weights: gender (2.1), age (0.06), stage (0.9), absence/presence of B-symptoms (A/B) (0.9), international prognostic index (IPI-NCCN) (1.1), body mass index (BMI) (-0.2), number of chemotherapy courses (0.9), Charlson comorbidity index (CCI) (0.3), bone mineral density after completion of chemotherapy (HU C) (-0.08), β-2-microglobulin (B2M) level (0.0007), lactate dehydrogenase (LDH) (0.006), body surface area (BSA) (-3.3). To assess the level of confidence in the proposed model for predicting bone density disorders in oncological pathology, ROC analysis was performed to obtain the corresponding curves and the area under them was estimated. A conclusion was made about the quality of the classification and the sensitivity, specificity, prognostic value of positive and negative results, the ratio of the probability of positive and negative results, as well as the accuracy of the classification were determined. For each of the four degrees of severity of violations (1C, 2C, 3C, 4C), it is necessary to carry out appropriate calculations, the matrices of inconsistencies for which are given in four tables. Sensitivity was calculated for 1C (98.8%), 2C (97.5%), 3C (95.2%) and 4C (98.5%); specificity for 1C (90.4%), 2C (83.3%), 3C (90.9%) and 4C (95%); predictive value of a positive result for 1C (97.6%), 2C (95.2%), 3C (97.5%) and 4C (97%); predictive value of a negative result for 1C (95%), 2C (90.9%), 3C (83.3%) and 4C (97.4%); accuracy for 1C (97.1%), 2C (97.1%), 3C (97.1%), and 4C (97.1%). According to the results of the analysis of ROC curves, a high level of classification of 1C (AUC=0.869), 3C (AUC=0.869) and 4C (AUC=0.869) was established. The average level of classification of bone density disorders according to 2C (AUC=0.758).

Keywords - Analytical analysis of publications, machine learning methods, forecasting, bone density, regression analysis, oncology, ROC analysis.

АБУБАКАР Садік Абдулхамід
Тернопільський національний технічний університет імені Івана Пулюя
СВЕРСТЮК Андрій
Тернопільський національний технічний університет імені Івана Пулюя,
Тернопільський національний медичний університет імені І. Я. Горбачевського МОЗ

АНАЛІТИЧНИЙ ОГЛЯД ПУБЛІКАЦІЙ ПО МЕТОДАХ МАШИННОГО НАВЧАННЯ В ОНКОЛОГІЇ ТА ПІДХІД ДО ОЦІНЮВАННЯ ЇХ ЯКОСТІ

У роботі проведено аналітичний огляд публікацій по методах машинного навчання в онкології та підхід до оцінювання їх якості. Проведено аналіз публікацій по роках в наукометричних базах Web of Science та Scopus. Представлено найбільшу кількість авторів, кількість публікацій серед університетів, кількість країн, категорій публікацій в наукометричній базі Scopus по методах машинного навчання в онкології. Запропонована багатофакторна регресійна модель прогнозування щільності кісткової тканини при онкологічній патології при прогнозуванні чотирьох ступенів важкості протікання досліджуваного захворювання. До даної моделі ввійшли такі фактори з відповідними ваговими коефіцієнтами: стать (2.1), вік (0.06), стадія (0.9), відсутність/наявність B-симптомів (A/B) (0.9), міжнародний прогностичний індекс (IPI-NCCN) (1.1), індекс маси тіла (BMI) (-0.2), кількість курсів хіміотерапії (Number of chemotherapy courses) (0.9), індекс коморбідності Шарльсона (ІКШ) (Charlson Comorbidity Index (CCI)) (0.3), мінеральна щільність кістки після завершення хіміотерапії (HU C) (-0.08), рівень β-2-мікроглобуліну (B2M) (0.0007), лактатдегідрогеназа (LDH) (0.006), площа поверхні тіла (BSA) (-3.3). Для оцінювання рівня довіри до запропонованої моделі прогнозування порушень щільності кісткової тканини при онкологічній патології проведено ROC-аналіз із отриманням відповідних кривих та оцінено площу під ними. Зроблено висновок про якість класифікації, а також визначено чутливість, специфічність, прогностичну цінність позитивного та негативного результатів, відношення правдоподібності позитивного та негативного результатів, а також точність класифікації. Відповідні розрахунки провести для кожного із чотирьох ступенів важкості порушень (1С, 2С, 3С, 4С), матриці невідповідностей до яких наведені у чотирьох таблицях. Розраховано чутливість для 1С (98,8%), 2С (97,5%), 3С (95,2%) та 4 С (98,5%); специфічність для 1С (90,4%), 2С (83,3%), 3С (90,9%) та 4 С (95%); прогностичну цінність позитивного результату для 1С (97,6%), 2С (95,2%), 3С (97,5%) та 4С (97%); прогностичну цінність негативного результату для 1С (95%), 2С (90,9%), 3С (83,3%) та 4С (97,4%.); точність для 1С (97,1%), 2С (97,1%), 3С (97,1%) та 4С (97,1%). За результатами аналізу ROC-кривих встановлено високий рівень класифікації 1С (AUC=0.869), 3С (AUC=0.869) та 4С (AUC=0.869). Середній рівень класифікації порушень щільності кісткової тканини для 2С (AUC=0.758).

Ключові слова: аналітичний аналіз публікацій, методи машинного навчання, прогнозування, щільність кісткової тканини, регресійний аналіз, онкологія, ROC-аналіз

Introduction

Accurate disease prediction is crucial in the medical field. Modern research demonstrates the effective combination of artificial intelligence algorithms and medical scientific research in healthcare . This article provides an analytical review of publications on machine learning methods in oncology and the approach to evaluating their quality. The choice of an artificial intelligence component for consultative and diagnostic information technology for glaucoma diagnosis [1] is highly important in modern medicine and technology. Integrating artificial intelligence into diagnostic technology can significantly improve the accuracy and accessibility of diagnosis, enabling early detection and personalized treatment of diseases, especially in regions with limited access to specialized medical professionals. This can lead to a substantial increase in effectiveness and improvement in patient health. Approach to developing an information system for monitoring patients with infectious diseases is discussed in [2]

Article [3] relates to CUDA-based parallelization of gradient boosting and bagging algorithms for disease diagnosis in the field of healthcare and technology.

In works [4-8], multifactor regression prediction models are built for diagnosis and risk assessment of various diseases. The CART algorithm - an acronym for Classification and Regression Trees - is often used. It is a decision tree algorithm that can be used for both classification and regression tasks. CART was developed by Leo Breiman, Jerome Friedman, Charles J. Stone, and Richard Olsen [9]. This algorithm can be divided into steps [10].

The tree is formed by optimizing the Gini index at each step. This approach allows building a tree that makes optimal splits for classification [11]. At each stage of tree construction, the optimal discriminant attribute that maximizes the Gini index or minimizes error is selected. Optimization methods such as searching for all possible splitting attributes and their thresholds can be used for this purpose. Through a recursive process, the selection of a separate attribute and its threshold for each new node is repeated until a stop condition is met.

Recursion ends when a stop condition is reached, such as the maximum tree depth, which in this case is set to 4. The goal of the CART algorithm is to minimize this Gini index at each split to obtain a more accurate and clean classification tree.

After building the model based on the decision tree, the next step is to analyze the accuracy and sensitivity of the model [12]. Sensitivity, also known as True Positive Rate (TPR) or Recall, is a metric that measures the ability of a classification model to correctly identify positive cases among the total number of actual positive cases [13]. This metric is critically important in healthcare, where failure to identify a condition (such as cancer or osteoporosis) correctly can have serious or even fatal consequences. Accuracy measures the percentage of correct identifications of positive cases. In healthcare, high accuracy means a more reliable diagnosis. Lower accuracy can lead to false positives, which can result in unnecessary treatment or tests, causing stress and additional expenses on medical services. To evaluate the model's quality, it is necessary to obtain the model's classification accuracy using Python [14].

In works [15 - 18], the effectiveness of the GINI index and information gain metrics in classification tasks using decision tree classifier algorithms is investigated. Decision trees are widely used in machine learning for classification and regression tasks due to their simplicity and interpretability. The GINI index and information gain are commonly employed to determine the best split at each node of the decision tree. The research aims to evaluate how these metrics influence the performance of decision tree classifiers in various classification tasks. By comparing their impact on accuracy, precision, recall, and F1 score metrics, this study provides insights into the strengths and limitations of using the GINI index and information gain in decision tree-based classification algorithms.

In works [19 - 21], the research on employing a decision tree-based approach, specifically the Classification and Regression Trees (CART) model, for diagnosing coronary artery disease (CAD). Coronary artery disease is a common and serious heart condition, and early diagnosis is crucial for effective treatment and management. Decision tree models, such as CART, are well-suited for medical diagnosis tasks due to their ability to handle complex decision-making processes and provide interpretable results. This study aims to develop a CART model trained on relevant medical data to accurately classify patients as either having or not having CAD based on their clinical parameters and test results. By evaluating the performance of the CART model in terms of sensitivity, specificity, accuracy, and other relevant metrics, this research contributes to the development of efficient and reliable diagnostic tools for CAD.

Machine learning using python is introduced to data scientists in works [22, 23], focusing on practical applications with Scikit-Learn and TensorFlow libraries.

The article [22], explores the potential applications of reinforcement learning in modern scenarios. It delves into how reinforcement learning algorithms can be utilised in various fields such as robotics, gaming, finance, and healthcare. By providing insights into the capabilities and limitations of reinforcement learning, the article highlights its significance in addressing complex decision-making tasks and optimising system performance in real-world settings.

Analytical review of publications on machine learning methods in oncology

To assess the relevance of research on machine learning methods and their use in oncology diagnostics in the Web of Science bibliometric database, an analytical query was formulated as follows:

(TS=("forecasting models") OR TS=("forecasting methods") OR TS=("PCA method") OR TS=("regression analysis") OR TS=("decision trees ") OR TS=("neural network models") OR TS=("cluster analysis") OR TS=("decision making methods") OR TS=("medical calculator") OR TS=("artificial intelligence") OR TS=("information system") OR TS=("expert system") OR TS=("reinforcement learning") OR TS=("regularisation") OR TS=("Markov decision process")) AND (TS=("oncology")).The search yielded 5046 publications, the distribution of which by year is shown in Figure 1.

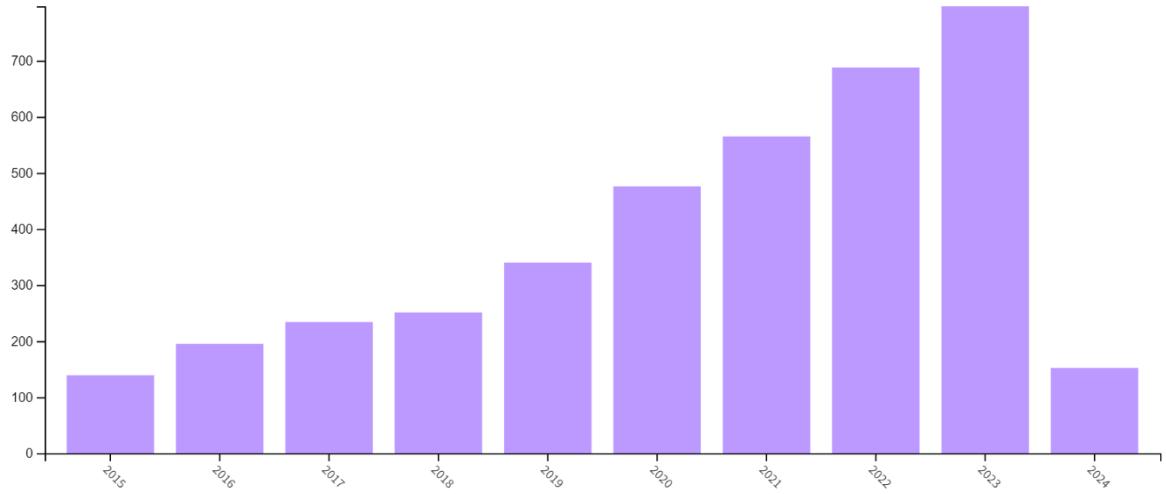

Fig. 1. Number of publications over the years in the Web of Science bibliometric database based on the search results of machine learning methods and their application in oncology diagnostics

According to Figure 1, over the past 10 years, there has been a steep increase in interest in the development and application of machine learning methods in oncology diagnostics. Specifically, in 2020, 639 publications were published, in 2021 – 771, in 2022 – 894, in 2023 – 1049, and initially in 2024 – 211.

Similarly, based on a search for machine learning methods and their application in oncology diagnostics in the Scopus bibliometric database, 538 publications were obtained. The distribution of these publications by year is shown in Figure 2.

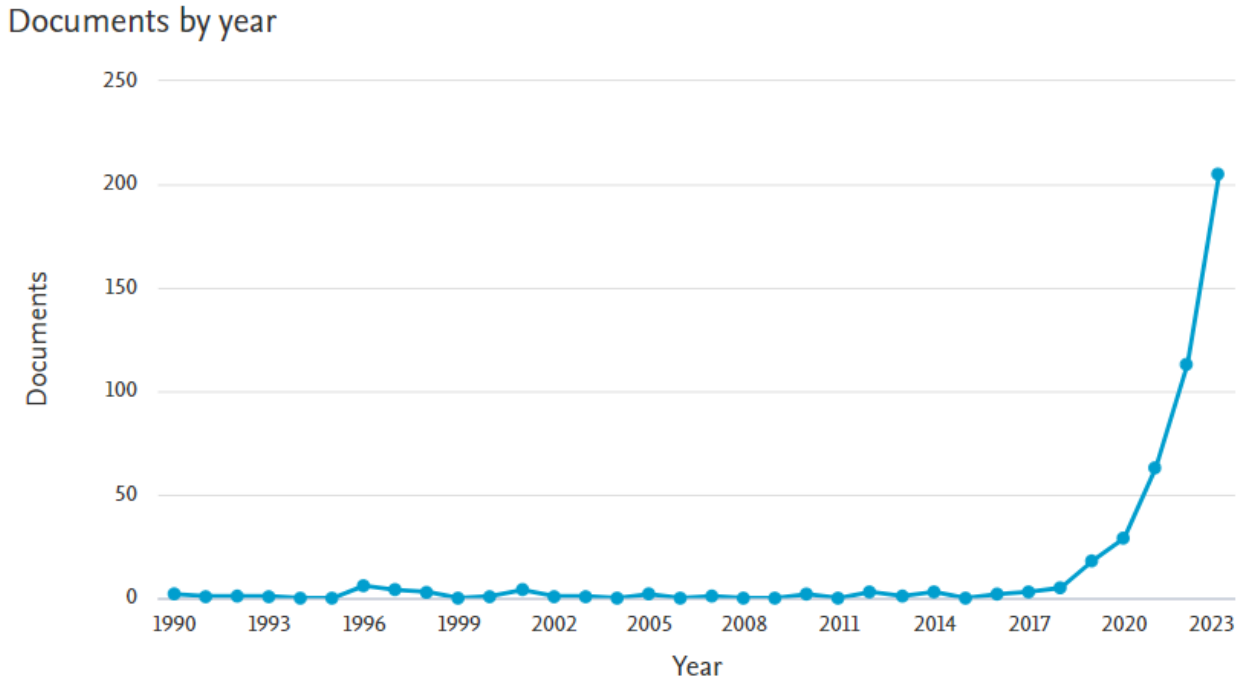

Fig. 2. Number of publications over the years in the Scopus bibliometric database based on the search results of machine learning methods and their application in oncology diagnostics

According to Figure 2, over the past 10 years, there has been a steep increase in interest in the development and application of machine learning methods in oncology diagnostics. Specifically, in 2020, 29 publications were published, in 2021 – 63, in 2022 – 113, in 2023 – 205, and initially in 2024 – 24.

Figure 3 shows the highest number of authors in the Scopus bibliometric database based on the search results of machine learning methods and their application in oncology diagnostics.

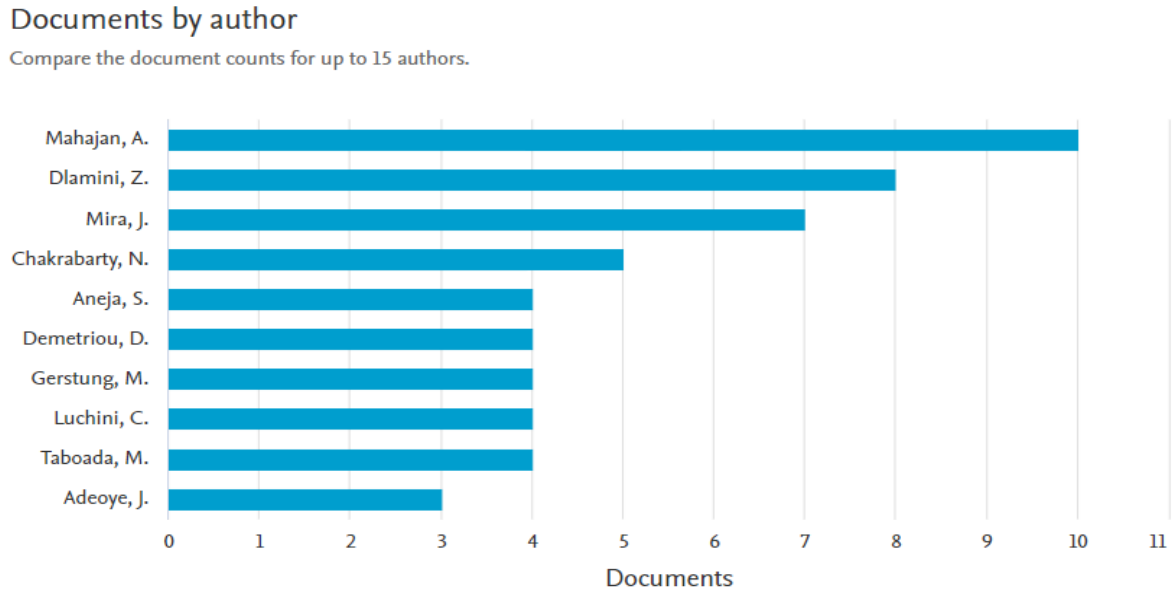

Fig. 3. Maximum number of authors in the Scopus bibliographic database based on the search results of machine learning methods and their application in oncology diagnostics

According to Figure 3, the number of documents by authors in the Scopus bibliography based on the search results of machine learning methods and their application in oncology diagnostics is as follows: Mahajan, A. – 10, Dlamini, Z. – 8, Mira, J. – 7, Chakrabarty, N. – 5, Aneja, S. – 4, Demetriou, D. – 4, Gerstung, M. – 4, Luchini, C. – 4, Taboada, M. – 4, Adeoye, J. – 3.

Figure 4 shows the highest number of publications among universities in the Scopus database based on the search results of machine learning methods and their application in oncology diagnostics.

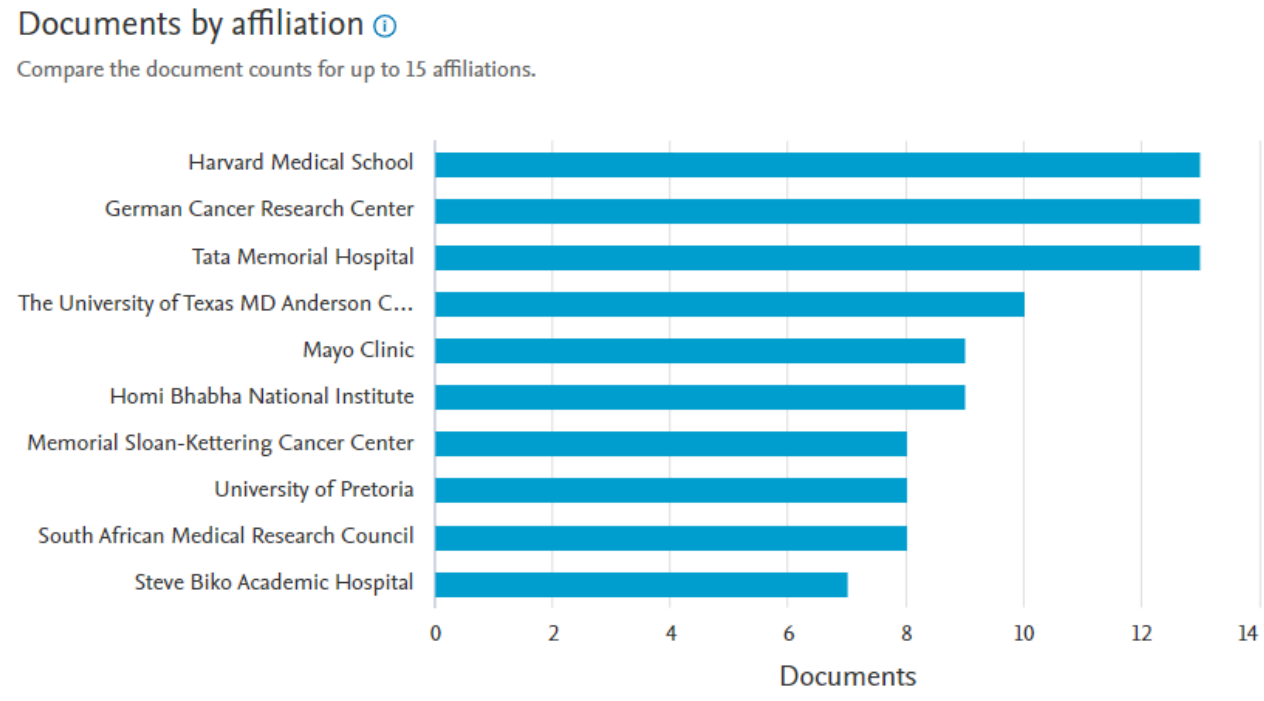

Fig. 4. Highest number of publications among universities in the Scopus bibliographic database based on the search results of machine learning methods and their application in oncology diagnostics

According to Figure 4, the number of affiliations in the Scopus bibliography based on the search results of machine learning methods and their application in oncology diagnostics are as follows: Harvard Medical School – 13, German Cancer Research Center – 13, Tata Memorial Hospital – 13, University of Texas – 10, Mayo Clinic – 9, Homi Bhabha National Institute – 8, Memorial Sloan Kettering Cancer Center – 8, University of Pretoria – 8, African Medical Research Council – 8, and Steve Biko Academic Hospital – 7.

Figure 5 shows the highest number of countries in the Scopus bibliographic database based on the search results of machine learning methods and their application in oncology diagnostics.

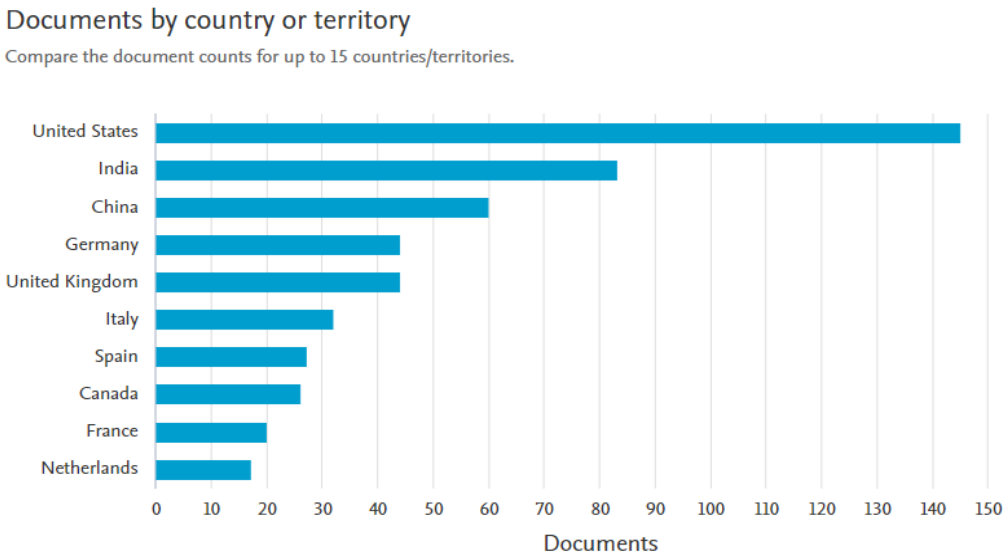

Fig. 5. The highest number of countries in the Scopus bibliographic database based on the search results of machine learning methods and their application in oncology diagnostics

Documents by countries and territories are presented in Figure 5. USA – 130, India – 83, China – 60, Germany – 44, United Kingdom – 44, Italy – 32, Spain – 28, Canada – 26, France – 20, Netherlands – 18.

Figure 6 shows the largest categories of publications in the Scopus bibliographic database based on the search results of machine learning methods and their application in oncology diagnostics.

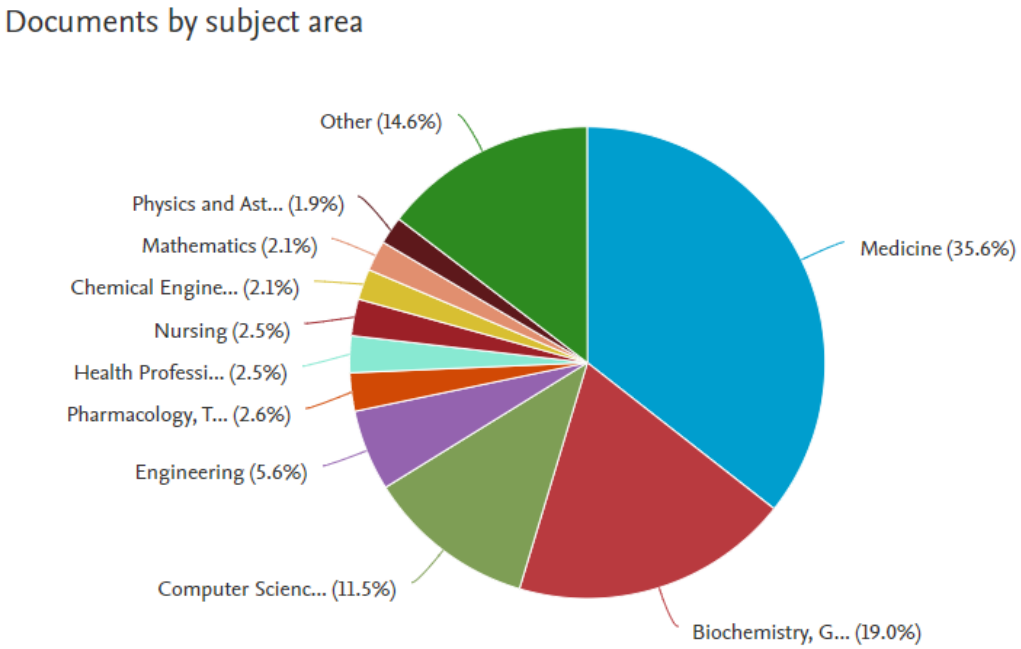

Fig. 6. Categories of publications in the Scopus bibliographic database based on the search results of machine learning methods and their application in oncology diagnostics

According to Figure 6, the documents are categorized as follows: Medicine - 35.6%, Biochemistry - 19.0%, Informatics - 11.5%, Engineering - 5.6%, Pharmacology - 2.6%, Healthcare - 2.5%, Nursing - 2.5%, Chemical Engineering - 2.1%, Mathematics - 2.1%, Physics and Astronomy - 1.9%.

Building a multi-factor regression model for predicting bone tissue density in oncological pathology (BTDOP)

A dataset consisting of 18 main factors was used to construct the multifactorial regression model for predicting BTDOP. The input data included age, gender, disease stage, presence or absence of B-symptoms, extracellular factors, molecular subtypes, international prognostic index, history of fractures, body mass index, number of chemotherapy courses, Charlson comorbidity index, bone mineral density at diagnostic stage, bone mineral density after chemotherapy completion, percentage decrease in bone mineral density after chemotherapy compared to diagnostic stage, levels of β-2 microglobulin, lactate dehydrogenase, body surface area, and changes in structural and functional status of bone tissue, risk degree. The aforementioned indicators of 115 patients with bone tissue density disorders were used to construct the multifactorial regression model for predicting bone tissue density in oncological pathology (BTDOP).

The results of obtaining significant factors in the Statistica 12.0 program and their weight coefficients (b) with corresponding levels of significance (p-value) in predicting BTDOP are presented in Figure 7.

N=115	b*	Std.Err. of b*	b	Std.Err. of b	t(102)	p-value
Intercept			34,53773	2,635856	13,1030	0,000000
Gender	0,112529	0,019220	2,13967	0,365463	5,8547	0,000000
Age	0,089086	0,032888	0,05533	0,020427	2,7088	0,007921
Stage	0,082780	0,034687	0,85176	0,356903	2,3865	0,018852
A/B	0,045375	0,017473	0,86541	0,333258	2,5968	0,010799
IPI-NCCN	0,214477	0,031357	1,09900	0,160677	6,8398	0,000000
BMI	-0,111279	0,035931	-0,23653	0,076374	-3,0970	0,002525
Number of chemotherapy courses	0,149445	0,027146	0,90325	0,164074	5,5052	0,000000
Charlson Comorbidity Index (CCI)	0,101023	0,032687	0,26975	0,087279	3,0907	0,002575
HU C	-0,404111	0,025615	-0,07865	0,004985	-15,7763	0,000000
B2M	0,078390	0,019605	0,00068	0,000169	3,9985	0,000121
LDH	0,081267	0,020085	0,00588	0,001452	4,0462	0,000101
BSA	-0,079969	0,039208	-3,29485	1,615443	-2,0396	0,043976

Fig. 7. The results of obtaining significant factors in the Statistica 12.0 program and their weight coefficients (b) with corresponding levels of significance (p-value) in predicting bone tissue density disorders (BTDOP)

According to Figure 7, the predictors included in the model for predicting BTDOP are gender (2.1), age (0.06), stage (0.9), absence/presence of B-symptoms (A/B) (0.9), international prognostic index (IPI-NCCN) (1.1), body mass index (BMI) (-0.2), number of chemotherapy courses (0.9), Charlson comorbidity index (CCI) (0.3), bone mineral density after completion of chemotherapy (HU C) (-0.08), β-2-microglobulin (B2M) level (0.0007), lactate dehydrogenase (LDH) (0.006), body surface area (BSA) (-3.3).

Evaluation of the model's performance in predicting bone tissue density disorders in oncology patients

To assess the level of confidence in the proposed model for predicting bone tissue density disorders (BTDD), it is necessary to conduct ROC analysis to obtain the corresponding curves and evaluate the area under the curves (AUC) to make conclusions about classification quality. Additionally, sensitivity, specificity, positive and negative predictive values, likelihood ratios of positive and negative results, as well as classification accuracy should be determined. The calculations should be performed for each of the four severity levels of BTDOP (1C, 2C, 3C, 4C), with confusion matrices provided in Tables 1, 2, 3, and 4.

Table 1

Number of sick patients to verify the prediction model of BTDOP classification for Stage 1 (1C)

Risk level	The number of patients with bone density disorders to verify the model of BTDOP for classification of 1C relative to 2C, 3C, and 4C is as follows:				
	True Positive 2C, 3C, 4C (a_{1234})	Sum 2C, 3C, 4C (a_{1234})	False Positive 1C (b_{1234})	Sum 1C (b_{1234})	Total (a_{1234}+b_{1234})
2C	20	83	2	2	85
3C	24		-		

4C	39		-		
1C	False Negative (c_{1234})	1	True Negative (d_{1234})	19	Total ($c_{1234}+d_{1234}$)
					20
Total:	$a_{1234}+c_{1234}$		$b_{1234}+d_{1234}$		$a_{1234}+b_{1234}+c_{1234}+d_{1234}$
	84		21		105

Given the numerical values from Table 1, we obtained:
Sensitivity of detecting BTDD 1C relative to 2C, 3C, and 4C:

$$Se_{1234}=(a_{1234}/(a_{1234}+c_{1234}))*100\% \quad (1)$$

$$Se_{1234}=(83/(83+1))*100\% = (83/84)*100\%=98,8\%.$$

Specificity of detecting changes in BTDOP: 1C relative to 2C, 3C, and 4C:

$$Sp_{1234}=(d_{1234}/(b_{1234}+d_{1234}))*100\% \quad (2)$$

$$Sp_{1234} = (19/(19+2))*100\% = (19/21)*100\%=90,4\%.$$

The positive predictive value of classifying patients with 1C relative to 2C, 3C, and 4C:

$$PPV_{1234}=(a_{1234}/(a_{1234}+b_{1234}))*100\% \quad (3)$$

$$PPV_{1234} =(83/(83+2))*100\%=(83/85)*100\%=97,6\%.$$

The negative predictive value of classifying patients with 1C relative to 2C, 3C, and 4C:

$$NPV_{1234}=(d_{1234}/ (c_{1234}+d_{1234}))*100\% \quad (4)$$

$$NPV_{1234} =(19/(19+1))*100\%=(19/20)*100\%=95\%.$$

The likelihood ratio of a positive result in detecting patients with 1C relative to 2C, 3C, and 4C:

$$LR+_{1234} = (Se_{1234}/(100-Sp_{1234})) \quad (5)$$

$$LR+_{1234} = (98,8/(100-90,4)) = 98,8/9,6=10,29$$

Similarly, the probability of obtaining a positive result of CTOP changes in patients with 2C, 3C, and 4C is 10.29 times higher compared to the probability of a positive result in patients with 1C.
The likelihood ratio of a negative result in detecting patients with 1C relative to 2C, 3C, and 4C:

$$LR-_{1234}=((100-Se_{1234})/Sp_{1234}) \quad (6)$$

$$LR-_{1234}=((100-98,8)/90,4)= 0,013$$

Therefore, the probability of obtaining a negative result of BTDOP changes in patients with 1C is 76.9 times higher (1/0.013) compared to the probability of a positive result in patients with 2C, 3C, and 4C.

$$\text{Accuracy } BTDOP_{1234} =((a_{1234}+d_{1234})/(a_{1234}+b_{1234} +c_{1234}+d_{1234}))*100\% \quad (7)$$

$$\text{Accuracy } BTDOP_{1234} =((83+19)/(83+2+1+19))*100\%=(102/105)*100\%=97,1\%$$

Therefore, the accuracy rate of predicting 1C is 97.1%.

Table 2

Number of diseased patients for validating the changes model of BTDOP classification for 2C

Risk level	The number of sick patients for verifying the model of BTDOP classification for 2C relative to 1C, 3C, 4C.				
	True Positive 1C, 3C, 4C (a_{2134})	Sum 1C, 3C, 4C (a_{2134})	False Positive 2C (b_{2134})	Sum 2C (b_{2134})	Total ($a_{2134}+b_{2134}$)
1C	18	79	2	4	83
3C	22		2		
4C	39		-		
2C	False Negative (c_{2134})	2	True Negative (d_{2134})	20	Total ($c_{2134}+d_{2134}$)
					22
Total:	$a_{2134}+c_{2134}$		$b_{2134}+d_{2134}$		$a_{2134}+b_{2134}+c_{2134}+d_{2134}$
	81		24		105

Given the numerical values from Table 2, we obtained:
Sensitivity of detecting BTDOP 2C relative to 1C, 3C, and 4C:

$$Se_{2134}=(a_{2134}/(a_{2134}+c_{2134}))*100\% \quad (8)$$

$$Se_{2134}= (79/(79+2))*100\% = (79/81)*100\%=97,5\%.$$

Specificity of detecting changes in BTDOP: 2C relative to 1C, 3C, and 4C:

$$Sp_{2134}=(d_{2134}/(b_{2134}+d_{2134}))*100\% \quad (9)$$

$$Sp_{2134} = (20/ (20+4)) *100\% = (20/24) *100\%=83,3\%.$$

The positive predictive value of classifying patients with 2C relative to 1C, 3C, and 4C:

$$PPV_{2134}=(a_{2134}/(a_{2134}+b_{2134}))*100\% \quad (10)$$

$$PPV_{2134} =(79/(79+4))*100\%=(79/83)*100\%=95,2\%.$$

The negative predictive value of classifying patients with 2C relative to 1C, 3C, and 4C:

$$NPV_{2134}=(d_{2134}/ (c_{2134}+d_{2134}))*100\% \quad (11)$$

$$NPV_{2134} =(20/ (20+2))*100\%=(20/22)*100\%=90,9\%.$$

The likelihood ratio of a positive result in detecting patients with 2C relative to 1C, 3C, and 4C:

$$LR+_{2134} = (Se_{2134}/ (100-Sp_{2134})) \quad (12)$$

$$LR+_{2134} = (97,5 / (100-83,3)) = 97,5 /16,7=5,83$$

Therefore, the probability of obtaining a positive result of CTOP changes in patients with 1C, 3C, and 4C is 5.83 times higher compared to the probability of a positive result in patients with 2C.

The likelihood ratio of a negative result in detecting patients with 2C, relative to 1C, 3C, and 4C:

$$LR-_{2134}=((100-Se_{2134})/ Sp_{2134}) \quad (13)$$

$$LR-_{2134}=((100-97,5)/ 83,3)= 0,03$$

Therefore, the probability of obtaining a negative result of CTOP changes in patients with 2C is 33.3 times higher (1/0.03) compared to the probability of a positive result in patients with 1C, 3C, and 4C.

$$\text{Accuracy } BTDOP_{2134} =((a_{2134}+d_{2134})/(a_{2134}+b_{2134} +c_{2134}+d_{2134}))*100\% \quad (14)$$

$$\text{Accuracy } BTDOP_{2134} =((79+20)/(79+2+2+20))*100\%=(99/103)*100\%=97,1\%$$

Therefore, the accuracy rate of predicting 2C is 96.1%.

Table 3

Number of diseased patients for validating the changes model of BTDOP classification for 3C

Risk level	The number of sick patients for verifying the model of BTDOP classification for 3C relative to 1C, 2C, 4C.				
	True Postive 1C, 2C, 4C (a_{3124})	Sum 1C, 2C, 4C (a_{3124})	False Negative 3C (b_{3124})	Sum 3C (b_{3124})	Total ($a_{3124}+b_{3124}$)
1C	20	79	-	2	81
2C	21		1		
4C	38		1		
3C	False Negative (c_{3124})	4	True Negative (d_{3124})	20	Total ($c_{3124}+d_{3124}$)
					24
Total:	$a_{3124}+c_{3124}$		$b_{3124}+d_{3124}$		$a_{3124}+b_{3124}+c_{3124}+d_{3124}$
	83		22		105

Given the numerical values from Table 3, we obtained:

Sensitivity of detecting BTDOP 3C relative to 1C, 2C, and 4C:

$$Se_{3124}=(a_{3124}/(a_{3124}+c_{3124}))*100\% \quad (15)$$

$$Se_{3124}= (79/(79+4)) * 100\% = (79/83) *100\%=95,2\%.$$

Specificity of detecting changes in BTDOP: 3C relative to 1C, 2C, and 4C:

$$Sp_{3124}=(d_{3124}/ (b_{3124}+d_{3124}))*100\% \quad (16)$$

$$Sp_{3124} = (20/ (20+2)) *100\% = (20/22) *100\%=90,9\%.$$

The positive predictive value of classifying patients with 3C relative to 1C, 2C, and 4C:

$$PPV_{3124}=(a_{3124}/ (a_{3124}+b_{3124}))*100\% \quad (17)$$

$$PPV_{3124} =(79 / (79+2))*100\%=(79/81)*100\%=97,5\%.$$

The negative predictive value of classifying patients with 3C relative to 1C, 2C, and 4C:

$$NPV_{3124}=(d_{3124}/ (c_{3124}+d_{3124}))*100\% \quad (18)$$

$$NPV_{3124} =(20/ (20+4))*100\%=(20/24)*100\%=83,3\%.$$

The likelihood ratio of a positive result in detecting patients with 2C relative to 1C, 3C, and 4C:

$$LR+_{3124} = (Se_{3124}/ (100-Sp_{3124})) \quad (19)$$

$$LR+_{3124} = (95,2/(100-90,9)) = 97,5 /9,1=10,7$$

Therefore, the probability of obtaining a positive result of CTOP changes in patients with 1C, 2C, and 4C is 10.7 times higher compared to the probability of a positive result in patients with 3C.

The likelihood ratio of a negative result in detecting patients with 3C, relative to 1C, 2C, and 4C:

$$LR\text{-}_{3124}=((100\text{-}Se_{3124})/ Sp_{3124}) \quad (20)$$

$$LR\text{-}_{3124}=((100\text{-}95,2)/ 90,9)= 0,05$$

Therefore, the probability of obtaining a negative result of CTOP changes in patients with 3C is 20 times higher (1/0.05) compared to the probability of a positive result in patients with 1C, 2C, and 4C.

$$\text{Accuracy BTDOP}_{3124} =((a_{3124}+d_{3124})/(a_{3124}+b_{3124} +c_{3124}+d_{3124}))*100\% \quad (21)$$

$$\text{Accuracy BTDOP}_{3124} =((79+20)/(79+2+4+20))*100\%=(99/105)*100\%=97,05\%$$

Therefore, the accuracy rate of predicting 3C is 97.05%.

Table 4

Number of diseased patients for validating the changes model of BTDOP classification for 4C

Risk level	The number of sick patients for verifying the model of BTDOP classification for 4C relative to 1C, 2C, 3C.				
	True Positive 1C, 2C, 3C (a_{4123})	Sum 1C, 2C, 3C (a_{4123})	False postive 4C (b_{4123})	Sum 4C (b_{4123})	Total ($a_{4123}+b_{4123}$)
1C	20	64	-	2	66
2C	22		-		
3C	22		2		
4C	False negative (c_{4123})	1	True negative (d_{4123})	38	Total ($c_{4123}+d_{4123}$)
					39
Total:	$a_{4123}+c_{4123}$		$b_{4123}+d_{4123}$		$a_{4123}+b_{4123}+c_{4123}+d_{4123}$
	65		40		105

Given the numerical values from Table 4, we obtained:

Sensitivity of detecting BTDOP 4C relative to 1C, 2C, and 3C:

$$Se_{4123}=(a_{4123}/(a_{4123}+c))*100\% \quad (22)$$

$$Se_{4123}= (64/(64+1))*100\% = (64/65)*100\%=98,5\%.$$

Specificity of detecting changes in BTDOP: 4C relative to 1C, 2C, and 3C:

$$Sp_{4123}=(d_{4123}/(b_{4123}+d_{4123}))*100\% \quad (23)$$

$$Sp_{4123}= (38/ (38+2)) *100\% = (38/40) *100\%=95\%.$$

The positive predictive value of classifying patients with 4C relative to 1C, 2C, and 3C:

$$PPV_{4123}=(a_{4123}/ (a_{4123}+b_{4123}))*100\% \quad (24)$$

$$PPV_{4123}=(64 / (64+2))*100\%=(64/66)*100\%=97\%.$$

The negative predictive value of classifying patients with 4C relative to 1C, 2C, and 3C:

$$NPV_{4123}=(d_{4123}/(c_{4123}+d_{4123}))*100\% \quad (25)$$

$$NPV_{4123}=(38/(38+1))*100\%=(38/39)*100\%=97,4\%.$$

The ratio of the likelihood of a positive result of detecting the use of 4C, deviation of 1C, 2C and 3C:

$$LR+_{4123} = (Se_{4123}/ (100\text{-}Sp_{4123})) \quad (26)$$

$$LR+_{4123} = (98,5/(100\text{-}95)) = 98,5 /5=19,7$$

Therefore, the probability of obtaining a positive result of BTDOP changes in patients with 1C, 2C, and 3C is 19.7 times higher compared to the probability of a positive result in patients with 4C.

The likelihood ratio of a negative result in detecting patients with 4C, relative to 1C, 2C, and 3C:

$$LR\text{-}_{4123}=((100\text{-}Se_{4123})/ Sp_{4123}) \quad (27)$$

$$LR\text{-}_{4123}=((100\text{-}98,5)/ 95)= 0,016$$

Therefore, the probability of obtaining a negative result of BTDOP changes in patients with 4C is 62.5 times higher (1/0.016) compared to the probability of a positive result in patients with 1C, 2C, and 3C.

$$\text{Accuracy BTDOP}_{4123} =((a_{4123}+d_{4123})/(a_{4123}+b_{4123} +c_{4123}+d_{4123}))*100\% \quad (28)$$

$$\text{Accuracy BTDOP}_{4123} =((64+38)/(64+2+1+38))*100\%=(102/105)*100\%=97,1\%$$

Therefore, the accuracy rate of predicting 4C is 97.14%.

ROC curve of the model predicting bone density disorders in oncology patients

ROC curves are often used to graphically display the relationship/tradeoff between clinical sensitivity and specificity for each possible cutoff value for a test or combination of tests. In addition, the area under the ROC curve gives an idea of the quality of the classification in question.

Since the area under the ROC curve is a measure of the classification quality overall, where a larger area means a more useful test, the areas under the ROC curves are used to compare the classification quality of different disease stages. The term ROC stands for receiver operating characteristic.

Figure 8 shows the ROC curve (relationship between sensitivity and specificity) of the model for predicting BTDOP disorders.

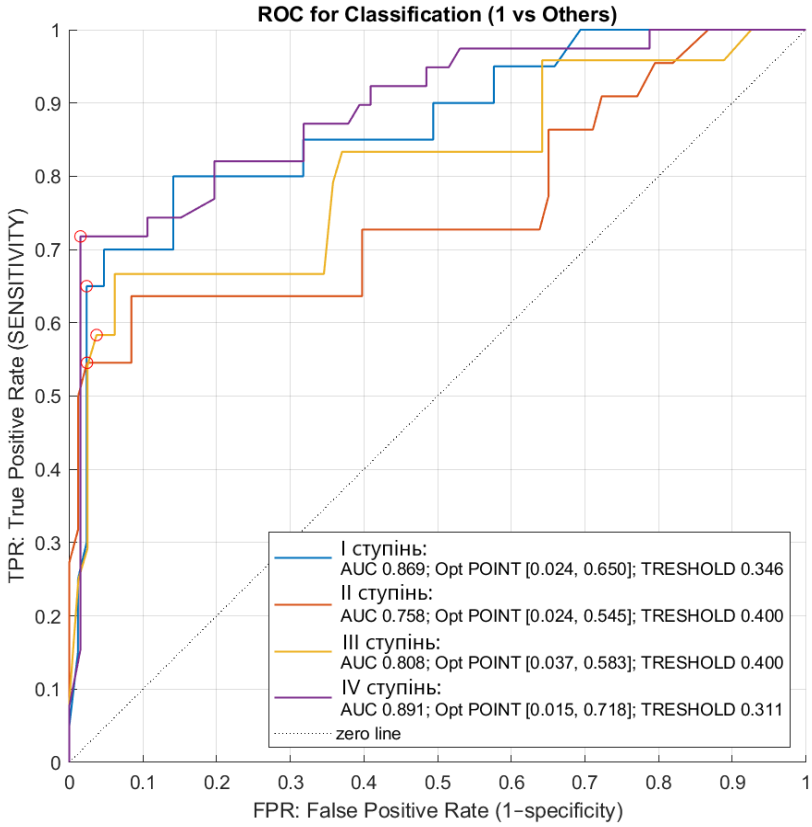

Fig. 8. ROC curve of the model predicting bone density disorders in oncology patients

Analysing Figure 8, it can be concluded that there is a high level of classification for grade 1 (AUC=0.869), grade 3 (AUC=0.869), and grade 4 (AUC=0.869) bone density disorders. The classification level for grade 2 bone density disorders is moderate (AUC=0.758).

Conclusion

This article provides an analytical review of publications on machine learning methods in oncology and an approach to evaluating their quality. The analysis of publications over the years in the Web of Science and Scopus databases is presented. The highest number of authors, publications among universities, the number of countries, and publication categories in the Scopus database on machine learning methods in oncology are presented. A multifactorial regression model for predicting bone mineral density disturbances in oncological pathology for predicting four severity levels of the studied disease course was proposed. The following factors were included in this model: gender (2.1), age (0.06), stage (0.85), absence/presence of B-symptoms (A/B) (0.87), international prognostic index (IPI-NCCN) (1.1), body mass index (BMI) (-0.24), number of chemotherapy courses (0.9), Charlson comorbidity index (CI) (0.27), bone mineral density after completion of chemotherapy (HU C) (-0.08), β-2-microglobulin (B2M) level (0.0007), lactate dehydrogenase (LDH) (0.006), body surface area (BSA) (-3.29). To assess the confidence level in the proposed model for predicting disturbances in bone mineral density in oncological pathology, ROC analysis was performed with obtaining corresponding curves and assessing the area under them. Conclusions were made regarding the classification quality, as well as sensitivity, specificity, prognostic value of positive and negative results, likelihood ratio of positive and negative results, and classification accuracy. The relevant calculations were performed for each of the four severity levels of disturbances (1C, 2C, 3C, 4C), with mismatch matrices provided in four tables. It was established that the high level of classification is for 1C (AUC=0.869), 3C (AUC=0.869), and 4C (AUC=0.869). The moderate level of classification of disturbances in bone mineral density in oncological pathology is for 2C (AUC=0.758).

References

1. Hovorushchenko, T., Kysil, V. (2023). Selection of the artificial intelligence component for consultative and diagnostic information technology for glaucoma diagnosis. In Computer systems and information technologies (Issue 4, pp. 87–90). Khmelnytskyi National University. https://doi.org/10.31891/csit-2023-4-12

2. Hovorushchenko, T., Zabelina, I., Rei, K., & Hovorushchenko, O. (2023). Method of creating an information system for monitoring infectious patients. In Computer systems and information technologies (Issue 3, pp. 59–64). Khmelnytskyi National University. https://doi.org/10.31891/csit-2023-3-7
3. Mochurad, L. (2023). Cuda-based parallelisation of gradient boosting and bagging algorithm for diagnosing diabetes. In Computer systems and information technologies (Issue 2, pp. 6–16). Khmelnytskyi National University. https://doi.org/10.31891/csit-2023-2-1
4. Herasymiuk M., Sverstiuk A., Kit I. Multifactor regression model for prediction of chronic rhinosinusitis recurrence. Wiadomości Lekarskie Medical Advances, Vol. LXXVI, Issue 5 Part I, May 2023. p. 928-935. DOI: 10.36740/WLek202305106.
5. Chukur O., Pasyechko N., Bob A., Sverstiuk A. Prediction of climacteric syndrome development in perimenopausal women with hypothyroidism (2022) Przeglad Menopauzalny, 21 (4), pp. 236 – 241. DOI: 10.5114/pm.2022.123522.
6. Musiienko V., Marushchak M., Sverstuik A., Filipyuk A., Krynytska I. Prediction Factors For The Risk Of Hypothyroidism Development In Type 2 Diabetic Patients. PharmacologyOnLine. 2021, Volume 3. p. 585-594.
7. Musiienko V., Sverstiuk A., Lepyavko A., Danchak, S., Lisnianska, N. Prediction factors for the risk of diffuse non-toxic goiter development in type 2 diabetic patients. Polski merkuriusz lekarski: organ Polskiego Towarzystwa Lekarskiegothis link is disabled, 2022, 50(296), pp. 94–98. PMID: 35436270.
8. Nykytyuk S.O., Sverstiuk A.S., Pyvovarchuk D.S., Klymnyuk S.I. A multifactorial model for predicting severe course and organ and systems damage in Lyme borreliosis in children. Modern pediatrics. No. 2(130) (2023). p. 6-16. DOI: https://doi.org/10.15574/SP.2023.130.6
9. Kaiser, H.F. An index of factorial simplicity. Psychometrika 39, 31–36 (1974). doi: 10.1007/bf02291575
10. Geron A. Applied machine-learning using Scikit-learn and TensorFlow. 2018. Vol. 1.
11. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. N Engl J Med. 1993 Sep 30;329(14):987-94. doi: 10.1056/NEJM199309303291402. PMID: 8141877.
12. Breiman, L.; Friedman, Jerome H. ; Olshen, Richard A. ; Stone, C. J.: Classification and Regression Trees. In: , 1984.
13. David James. 2018. Introduction to Machine Learning with Python: A Guide for Beginners in Data Science (1Ct. ed.). CreateSpace Independent Publishing Platform, North Charleston, SC, USA.
14. P.-N. Tan, M. Steinbach, and V. Kumar, “Introduction to Data Mining,” Pearson, 2005.
15. Tangirala, Suryakanthi. "Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm." International Journal of Advanced Computer Science and Applications 11.2 (2020): 612-619.
16. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. BMJ. 1994 Jun 11;308(6943):1552. doi: 10.1136/bmj.308.6943.1552.
17. Unpingco, Jose. Python for Probability, Statistics, and Machine Learning. 2019.
18. Leon Derczynski. Complementarity, F-score, and NLP Evaluation. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), Portorož, Slovenia. European Language Resources Association (ELRA). 2016.
19. Ghiasi, Mohammad M., Sohrab Zendehboudi, and Ali Asghar Mohsenipour. "Decision tree-based diagnosis of coronary artery disease: CART model." Computer methods and programs in biomedicine 192 (2020): 105400.
20. Martsenyuk V., Klos-Witkowska A., Sverstiuk A. Stability Investigation of Biosensor Model Based on Finite Lattice Difference Equations (2020) Springer Proceedings in Mathematics and Statistics, 312, pp. 297 – 321. DOI: 10.1007/978-3-030-35502-9_13.
21. Zhi Cheng, Minoru Nakatsugawa, et al. “Evaluation of classification and regression tree (CART) model in weight loss prediction following head and neck cancer radiation,” National Library of Medicine, 2017.
22. A.S Abdulhameed, S. Lupenko, “Potentials of Reinforcement Learning in Contemporary Scenarios,” Visnik TNTU, 2022.

Abubakar Sadiq Abdulhameed **Абубакар Садік Абдулхамід**	postgraduate student of the Faculty of Computer Sciences, Ternopil National Technical University named after Ivan Pulyu, Ternopil, Ukraine https://orcid.org/0000-0003-1867-4698 sadiq.ai@icloud.com	аспірант факультету комп'ютерних наук, Тернопільський національний технічний університет імені Івана Пулюя, Тернопіль, Україна.
Andriy Sverstyuk **Андрій Сверстюк**	professor, Doctor of Technical Sciences, Department of Medical Informatics, Ternopil National Medical University named after I. Ya. Gorbachevskii, Ministry of Health of Ukraine, Ternopil, Ukraine. Professor of the Department of Computer Sciences, Ternopil National Technical University named after Ivan Pulyu, Ternopil, Ukraine. https://orcid.org/0000-0001-8644-0776 sverstyuk@tdmu.edu.ua	професор, доктор технічних наук, кафедра медичної інформатики, Тернопільський національний медичний університет імені І. Я. Горбачевського МОЗ України, Тернопіль, Україна. Професор кафедри комп'ютерних наук, Тернопільський національний технічний університет імені Івана Пулюя, Тернопіль, Україна.