Journal of Experimental and Clinical Medicine
https://dergipark.org.tr/omujecm

Research Article

J Exp Clin Med
2023; 40(3): 633-639
doi: 10.52142/omujecm.40.3.35

Is there a correlation between dominant extremity and cervical disc herniation using machine learning methods?

Engin YÜCEL [1], Yener AKYUVA [2]*

[1]Department of Neurosurgery, Alanya Alaaddin Keykubat University Education and Research Hospital, Antalya, Türkiye
[2]Department of Neurosurgery, Tayfur Ata Sökmen Medical Faculty, Hatay Mustafa Kemal University, Hatay, Türkiye

Received: 18.08.2023 • **Accepted/Published Online:** 23.08.2023 • **Final Version:** 30.09.2023

Abstract

We aimed to investigate the possibility of determining the relationship between the dominant extremity and cervical disc herniation using a machine learning approach. A total of 561 patients diagnosed with cervical disc herniation were examined for dominant extremity, level and side of cervical disc herniation, and the nature of the herniation (calcified/soft). These patients formed the basis for a two-step machine learning system creation. The first step (included the data of 80% of the patients) focused on determining the type of cervical disc herniation by selecting the top five performing classification models out of 15 different models and tuning the hyperparameters. In the second step, the machine learning system was validated using data from a randomly selected subset of patients (20% of the patients). The study results showed that while most models performed well, the gradient boosting classifier was the most accurate (89.38%) for determining the herniated disc nature. However, for classifying the disc herniation direction, the models did not exhibit strong performance. Thus, machine learning can accurately identify the relationship between cervical disc herniation and dominant extremity with a high degree of accuracy.

Keywords: machine learning, artificial intelligence, disc degeneration, spinal disease

1. Introduction

Machine learning (ML) is rapidly being established in every aspect of modern life. Its ability to simultaneously process larger amounts of data and quickly perform multiple comparisons make it an integral part of our future (1). In the medical field, research and studies on ML are becoming increasingly prevalent. Due to the awareness generated by these efforts, the widespread adoption of ML in areas such as healthcare data systems and medical education is inevitable (2).

Because the cervical vertebrae bears a relatively lesser load than the lumbar vertebrae, experiences fewer traumas, and is less affected by environmental factors, they can be evaluated more easily using ML methods. The nature and determination of the direction of cervical disc herniation (CDH) has been a relatively underexplored area in artificial intelligence (AI) studies. To shed light on future studies, we aimed to approach this issue by incorporating AI (3).

2. Material and Methods

Between January 2020 and June 2023, a total of 561 patients who were diagnosed with CDH and had presented at the neurosurgery clinic of a tertiary university hospital were included in the study. The present study was performed in accordance with the framework of the Declaration of Helsinki and approved by the Alanya Alaaddin Keykubat University Ethics Committee (No: 11/06; approval date: 14.06.2023). The dominant extremity, level and side of CDH, and nature analysis of the radiological data (calcified (Fig. 1) or soft (Fig. 2), were documented for each patient. From these patients, 80% were randomly selected. The data of these selected patients were integrated into an AI module to create the main data table. Subsequently, the accuracy of different methods in generating data was investigated using the data of the remaining 20% of the patients. The suitability of the system was also explored in this context.

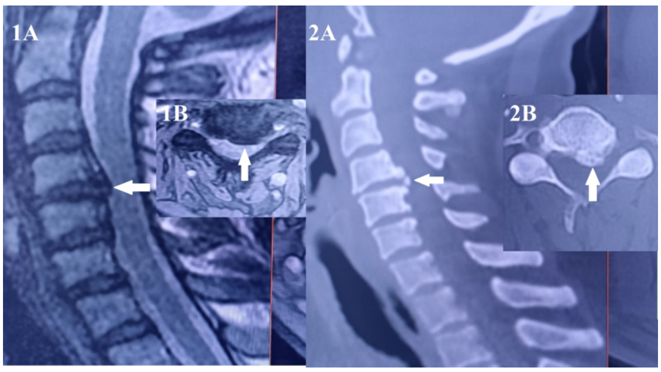

Fig. 1. Computed tomography (1 A, B) and magnetic resonance imaging (2 A, B) of CDH calcification

*Correspondence: yenerakyuva@hotmail.com

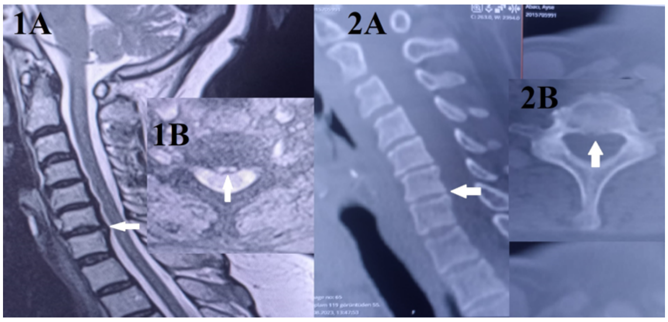

Fig. 2. Magnetic resonance imaging (1 A and B) and computed tomography (1 A and B) images of soft herniated disc

2.1. Statistical analysis

The normally distributed continuous variables are expressed as mean ± Standard Deviation (SD) and the non-normally distributed continuous variables are expressed as median (min-max). The Shapiro–Wilk test was used to determine the normality of the data. The categorical variables are expressed as frequencies (n) and percentages (%). The Pearson chi-square and Fisher's exact tests were used to determine the relationship between the categorical variables. The Kruskal–Wallis test was used for non-parametric comparisons of continuous data, and the independent t-test and one-way ANOVA were used for parametric comparisons. A post-hoc analysis was performed using the Bonferroni correction. All statistical analyses were carried out using IBM SPSS Statistics for Windows (version 23.0; IBM Corp., Armonk, NY). A two-sided $p < 0.05$ was considered statistically significant.

2.2. Methodology of machine learning

To be able to use classification models for determining the direction and nature of disc herniation, a custom dataset was created by collecting information from 561 patients treated clinically and surgically. The dataset included the following data: age, gender, dominant hand, CDH level and its direction and nature, and direction of arm pain (4).

Using this dataset, 15 different classification models were evaluated. The top five models that performed the best for each of the two outcomes (direction and nature) were selected, and their hyperparameters were tuned accordingly. Subsequently, the models were evaluated using metrics such as accuracy, precision, specificity, recall, F1-score, negative predictive value (NPV), and false positive rate (FPR).

The models used for classifying the nature of CDH were decision tree (DT) classifier, random forest classifier, gradient boosting (GB) classifier, multi-layer perceptron (MLP) classifier, and eXtreme gradient boosting (XGB) classifier. The models that produced the best results for determining the CDH direction were K-nearest neighbors (KNN) classifier, GB classifier, DT classifier, random forest classifier, and XGB classifier.

2.3. Models used

GB classifier

Boosting algorithms progressively combine weak learners, which perform marginally better than random guessing, to create strong learners. GB is a regression technique that shares similarities with boosting. It determines an estimate of the function that maps input samples to their corresponding output values by minimizing the error function's estimated value using the training data.

DT classifier

DTs are supervised learning models capable of handling classification and regression tasks, although they are predominantly used for solving classification problems. They have a tree-like structure, where each node corresponds to a feature value check, branches represent test outcomes, and leaf nodes represent the final classifications. DTs can efficiently generate interpretable rules and classify data with minimal computation.

Random forest classifier

The random forest classifier is a popular ML technique that leverages multiple DTs built on various subsets of the main dataset to make predictions. It functions as both a regression and classification model. As a regression model, it computes the mean of all the DT outcomes. As a classification model, it combines the votes from multiple DTs to obtain the final prediction.

KNN classifier

The KNN classifier prediction algorithm follows a lazy learning technique, generating predictions based on the KNN input. When predictions for any instance are requested, the entire prediction process is carried out. The Euclidean distance method is commonly used to determine the proximity between instances.

MLP classifier

The MLP classifier is a prediction algorithm based on an artificial neural network (MLP). When predictions for any instance are required, the neural network processes the input through its layers to generate the output. The MLP classifier is particularly effective for solving complex classification problems and can determine both linear and non-linear relationships in the data.

XGB classifier

The XGB classifier is a speedy and robust implementation of GB used for classification purposes. It leverages decision trees to make precise predictions, and the final outcome is determined through the combined voting of multiple trees.

2.4. Metrics used

Classification accuracy

The overall accuracy of the classifiers indicates the percentage of correct predictions among all predictions. The accuracy was calculated using the following equation: $Accuracy = (TP + TN) / (TP + TN + FP + FN)$, where TP is True Positives, TN is True Negatives, FP is False Positives, and FN is False

Negatives.

Precision

Precision is a crucial metric used to assess the classifier performance. It represents the ratio of true positives to the sum of true positives and false positives. The precision was calculated using the follow equation: $Precision = (TP) / (TP + FP)$, where TP is true positives and FP is false positives.

Recall/Sensitivity/True positive rate

True positive recall, commonly referred to as recall, is a metric defined as the ratio of true positive results to the sum of true positive and false negative results. Recall was calculated using the following equation: $Sensitivity = (TP) / (TP + FN)$, where TP is true positives and FN is false negatives.

FPR

The FPR is the ratio of false positive values to the sum of false positive values and true negative values. FPR was calculated using the following equation: $FPR = (FP) / (FP + TN)$, where FP is false positives and TN is true negatives.

NPV

The NPV is another significant metric used to assess the classifier performance. It represents the ratio of true negative values to the sum of true negative and false negative values. NPV was calculated using the following equation: $NPV = (TN) / (TN + FN)$, where TN is true negatives and FN is false negatives.

F1-score

The F-measure, also known as the F1-score, is determined by obtaining the harmonic mean of accuracy and recall. A value of 0 indicates the worst performance, while a value of 1 indicates the best performance. The F1-score was calculated using the following formula: $F1\text{-}score = 2 * precision * recall / (precision + recall)$

2.5. Experimental setup

All ML algorithms and classification and regression models used in this study were evaluated using the same dataset with the same split ratios in Google Colab (Global Al Hub, Matterhorn, Switzerland). The top five models that yielded the best results were selected, and their hyperparameters were tuned. Google Colab is a free platform that offers various tools for working with AI, allowing us to utilize ML effectively.

While deep learning and clustering models were utilized, their outcomes were not assessed due to their lower performance compared to the described classification models.

2.6. Machine learning dataset

This was a novel study we conducted in this field of study. Hence, there was no existing dataset to be used. To create our own custom dataset, we collected the following information from the 561 admitted symptomatic patients after obtaining their consent: age, gender, dominant hand, CDH level and its direction and nature, and direction of arm pain. The following eight features were used as inputs: age, gender, dominant hand, pain direction and CDH at C3-4, C4-5, C5-6, and C6-7). The remaining factors, herniation nature and CDH direction, were used as outputs.

Within our dataset, 443 patients had a soft disc and 118 patients had a hard disc. A total of 197 patients had a right-sided CDH, 222 patients had a left-sided CHD, and 142 patients had bilateral CDH. The training set comprised of 80% of the dataset, while the test set comprised of 20% of the dataset (Table 1).

Table 1. Detailed information about the dataset

No	Attribute Name	Abbreviation	Values	Explonation
1	Gender	Gender	0-1	Male-Female
2	Age	Age	21-81	21-81
3	Dominant Hand	dHand	0-2	Right-Left-Both
4	CDH Level C3-4	hLevel_C3-4	0-1	Yes-No
5	CDH Level C4-5	hLevel_C4-5	0-1	Yes-No
6	CDH Level C5-6	hLevel_C5-6	0-1	Yes-No
7	CDH Level C6-7	hLevel_C6-7	0-1	Yes-No
8	CDH Direction	hDirection	0-2	Right-Left-Both
9	Hernia Type	hType	0-1	Soft-Hard
10	Pain Direction	pDirection	0-2	Right-Left-Both

3. Results

The mean age of the patients in the outpatient clinic and surgical groups (48.97 ± 11.69 and 48.98 ± 10.15; $p = 0.995$) and the gender distributions ($p = 0.875$) were statistically similar. No significant differences were observed in the dominant hand ($p = 0.639$), CDH level ($p = 0.792$), and CDH nature ($p = 0.871$) between the two groups. A left-sided CDH was more commonly seen than a right-sided CDH in the outpatient clinic (40.9% vs. 26%). In the surgical group, a central CDH was seem more commonly than a right- or left-sided CDH (38% vs. 24.1%; $p = 0.048$). The proportion of patients with bilateral arm pain was higher in the surgical group than in the outpatient clinic group (24% vs. 6.7%) ($p = 0.001$) (Table 2).

Independent t-test, Pearson chi-square test, Fisher's exact test. Statistical analysis indicated no significant differences between the groups within the same column of lowercase letters.

Table 2. Demographics information of patients'

Variables	Clinic patients (n=511)	Surgical patients (n=50)	
Age (year)			
Mean±SD	48.97±11.69	48.98±10.15	0.995
Min-maks	21-81	33-73	
Gender, *n(%)*			
Male	200(39.1)	19(38.0)	0.875
Female	311(60.9)	31(62)	
Hand preference, *n(%)*			
Right	394(77.1)	37(74.0)	0.639
Left	50(9.8)	7(14.0)	
Both hands	67(13.1)	6(12.0)	
Level of CDH. *n(%)*			
C3-4	9(1.8)	1(2.0)	0.792
C4-5	59(11.5)	8(16.0)	
C5-6	284(55.8)	25(50.0)	
C6-7	159(31.1)	16(32.0)	
Direction of CDH, *n(%)*			
Right	179(35) [a]	18(36)[a]	**0.048**
Left	209(40.9) [b]	13(26) [a]	
Central	123(24.1) [b]	19(38)[a]	
Nature of CDH, *n(%)*			
Soft	402(79)	39(78)	0.871
Hard	107(21)	11(22)	
Direction of Arm Pain, *n(%)*			
Right	265(51.9)[a]	20(40)[a]	**0.001**
Left	212(41.5)[a]	18(36)[a]	
Both Sides	34(6.7)[b]	12(24)[a]	

Table 3. Patients' ages according to disease-related characteristics

	age			
Characteristics	**X̄±SD**	**Min-Maks**	**p**	**Post-Hoc (Adj P)**
Dominant hand				
Right	49.35±11.10	21-81	0.369	
Left	46.98±13.5	27-74		
Both Hands	48.26±12.85	29-75		
Level of CDH				
1.C3-4 (low count)	50.90±10.78	39-66	**<0.001**	1>2 (NS) 1>3(NS)
2.C4-5	45.03±10.31	21-68		1<4(NS) 2<3(NS)
3.C5-6	47.55±10.31	22-75		2<4 (<0.001) 3<4(<0.001)
4.C6-7	52.88±13.04	23-81		
Direction of CDH				
1. Right	48.56±11.64	22-81	**<0.001**	**1>2 (NS)**
2. Left	46.55±11.93	21-75		2<3(<0.001)
3.Central	53.32±11.24	27-76		1<3(0.001)
Nature of CDH				
Soft	45.55±9.25	21-68	**<0.001**	
Hard	61.60±10.55	28-81		
Direction of Arm Pain				
Right	48.50±10.72	22-77	0.147	
Left	48.95±12.69	21-81		
Both Sides	51.98±10.28	34-73		

The dominant hand (p = 0.369) and direction of arm pain (p = 0.147) were not associated with the patient's age. The mean age of the patients with CDH at the C6-7 level was higher than those with CDH at the C4-5 and C5-6 levels (p < 0.001). Patients with a central CDH or hard disc had higher mean ages than those with right- or left-sided CDH or soft disc (p < 0.001) (Table 3).

Independent t-test, one-way ANOVA, Kruskal–Wallis test. The same lowercase letters within a column indicate no significant difference between the groups. Bonferroni, NS: non-significant

No significant differences were observed in the distribution of dominant hand and direction of arm pain across age groups (p = 0.228). CDH was more prevalent at the C5-6 level in patients aged < 50 years than in those aged ≥ 50 years (60.1% vs. 48.6%). CDH was more prevalent at the C6-7 level in patients aged ≥ 50 than in those aged < 50 years (39.5% vs. 25.1%; p = 0.001). A left-sided CDH was more common in patients aged < 50 years old than in those aged ≥ 50 years (46.5% vs. 30.5%). A central CDH was more common in patients aged ≥ 50 than in those aged < 50 years (35.8% vs. 17.3%; p < 0.001). Furthermore, the prevalence of a hard disc was significantly higher in patients aged ≥ 50 than in those aged < 50 years (42.7% vs. 4.7%; p < 0.001) (Table 4).

Pearson chi-square test, Fisher's exact test. The same lowercase letters within a row indicate no significant difference between the groups.

There is no significant difference in the CDH level (p = 0.156), herniation direction (p = 0.095), and CDH nature (p = 0.318) according to the patient's dominant hand. Left-sided arm pain was higher in patients predominantly using the left hand than in those using both hands (54.4% vs. 32.9%). In the group using both hands, pain occurred more frequent on both sides than in the group predominantly using the left hand

(24.7% vs. 1.8%, p < 0.001). In the group using both hands, more patients had right-sided dominance than left-sided dominance (24.7% vs. 6.3%) (Table 5).

Table 4. Disease-Related Characteristics According to Age Groups of Patients

	Age		
Characteristics, n(%)	**<50 (n=318)**	**≥50 (n=243)**	**p**
Dominant hand			
Right	236(74.2)	195(80.2)	0.228
Left	37(11.6)	20(8.2)	
Both Hands	45(14.2)	28(11.5)	
Level of CDH			
C3-4	4(1.3)[a]	6(2.5)[a]	**0.001**
C4-5	44(13.8)[a]	23(9.5)[a]	
C5-6	191(60.1)[a]	118(48.6)[b]	
C6-7	79(25.1)[a]	96(39.5)[b]	
Direction of CDH			
Right	115(36.2)[a]	82(33.7)[a]	**<0.001**
Left	148(46.5)[a]	74(30.5)[b]	
Central	55(17.3)[a]	87(35.8)[b]	
Nature of CDH			
Soft	303(95.3)[a]	138(57.3)[b]	**<0.001**
Hard	15(4.7)[a]	103(42.7)[b]	
Direction of Arm Pain			
Right	160(50.3)	125(51.4)	0.213
Left	137(43.1)	93(38.3)	
Both Sides	21(6.6)	25(10.3)	

Pearson chi-square test, Fisher's exact test. The same lowercase letters within a row indicate no significant difference between the groups.

3.1. Machine learning analysis

The experiment used the herniation nature and direction data, which was split into the training (80%) and testing (20%) datasets. A confusion matrix measured the performance of the classifier models (Tables 2 and 4). The results of the study classifiers are listed in Tables 3 and 5.

Table 5. Other Characteristics According to Patients' Dominant Hand

	Dominant hand			
Characteristics, n(%)	**Right hand**	**Left hand**	**Both hands**	**p**
Level of CDH				
C3-4	10(2.3)	0(0)	0(0)	0.156
C4-5	57(12.3)	6(10.5)	4(5.5)	
C5-6	239(55.5)	29(50.9)	41(56.2)	
C6-7	125(29.0)	22(38.6)	28(38.4)	
Direction of CDH				
Right	156(3.2)	13(22.8)	28(38.4)	0.095
Left	162(37.6)	32(56.1)	28(38.4)	
Central	113(26.2)	12(21.1)	17(23.3)	
Nature of CDH				
Soft	337(78.6)	49(86.0)	55(75.3)	0.318
Hard	92(21.4)	8(14.0)	18(24.7)	
Direction of Arm Pain				
Right	229(53.1)[a]	25(43.9)[a]	31(42.5)[a]	**<0.001**
Left	175(40.6)[a.b]	31(54.4)[b]	24(32.9)[a]	
Both Hands	27(6.3)[a]	1(1.8)[a]	18(24.7)[b]	

The determination of herniation nature yielded the most successful classification model (Table 3), with the GB classifier achieving a performance of 89.3%. Although the other models were not as successful as GB, their performances were still quite impressive. The DT and random forest classifiers both achieved the same percentage result (87.6%), the XGB classifier achieved a performance of 86.7%, and the MLP classifier yielded a performance of 85.8% (Table 3).

The models of herniation direction determination did not achieve the same level of success as that of herniation nature classification models. The classification models could not establish sufficient correlations between the data or adequately generalize data. Thus, the RF was the most successful (62.8%), followed closely by GB (59.2%). The XGB, MLP, and KNN models achieved results of 52.2%, 51.3%, and 50.4%, respectively (Tables 6, 7, 8, and 9)

Table 6. Summarised Depiction of Confusion Matrices for Herniation Nature for All Classifiers

	Decision Tree Classifier	Random Forest Classifier	Gradient Boosting Classifier	MLP Classifier	XGB Classifier
TP	17	16	16	19	17
TN	81	81	82	67	81
FP	3	3	2	17	3
FN	12	13	13	10	12

Three different results were obtained regarding herniation direction. This is because the CDH could occur on the right, left, or bilaterally within our dataset. The models we used provided separate results for each scenario (Table 3).

The F1-score (Table 9) provided us with an idea of the model's classification capability; the higher the value, the higher the model's classification ability. The classification model with the highest accuracy also had the highest F1-score. For herniation nature determination, the GB classification model yielded the highest F1-score (0.854), while the MLP model yielded the lowest F1-score (0.788).

Table 7. Performance Statistics of All Classifiers for Herniation Nature

Model	Accuracy (%)	Precision	Sensitivity	Recall	F1-score	NPV (%)	FPR (%)
Decision Tree Classifier	87.6	0.883	0.894	0.781	0.814	87.2	2.3
Random Forest Classifier	87.6	0.883	0.894	0.781	0.814	87.2	2.3
Gradient Boost	89.3	0.898	0.904	0.815	0.845	89.1	2.3

Classifier							
MLP Classifier	85.8	0.851	0.842	0.758	0.788	86.1	3.5
XGB Classifier	86.7	0.860	0.850	0.775	0.804	87.0	3.5

Table 8. Summarised Depiction of Confusion Matrices for Herniation Direction for All Classifiers

	K-Neighbors Classifier	Random Forest Classifier	Gradient Boosting Classifier	MLP Classifier	XGB Classifier
TN-Right	49	50	50	69	51
TN-Left	53	57	58	0	54
TN-Both	68	74	71	83	67
FP- Right	20	19	19	0	18
FP- Left	21	17	16	74	20
FP- Both	15	9	12	0	16
TP- Right	25	32	30	0	24
TP- Left	24	26	26	39	25
TP- Both	8	10	10	0	10
FN- Right	19	12	14	44	20
FN- Left	15	13	13	0	14
FN- Both	22	20	20	30	20
Sensitivity- Right	0.555	62.7	61.2	Nan	57.1
Sensitivity- Left	0.533	60.4	61.9	34.5	55.5
Sensitivity- Both	0.347	52.6	45.4	Nan	34.8
NPV- Right (%)	72.0	80.6	78.1	61.0	71.8
NPV- Left (%)	77.9	81.4	81.6	Nan	79.4
NPV- Both (%)	77.5	78.7	78.0	74.3	77.0
FPR- Right (%)	28.9	27.5	27.5	0	26.0
FPR- Left (%)	28.3	22.9	21.6	1	27.0
FPR- Both (%)	18.0	10.8	14.4	0	19.2

Table 9. Performance statistics of All Classifiers for Herniation Direction

Model	Accuracy (%)	Precision	Recall	F1-score
K-Neighbors Classifier	50.4	0.478	0.483	0.478
Random Forest Classifier	62.8	0.612	0.602	0.599
Gradient Boosting Classifier	59.2	0.580	0.571	0.570
MLP Classifier	51.3	0.466	0.473	0.443
XGB Classifier	52.2	0.503	0.506	0.503

The NPV and FPR indicate the accuracy of detecting negative cases and the rate of incorrectly detecting negative values, respectively. Thus, a higher NPV and lower FPR indicate better results. Sensitivity shows how successful we are in achieving a true positive rate in correct outcomes. Finally, the recall value is necessary for calculating some key metrics, as mentioned before.

Thus, our classification models exhibit a performance of 89.3% for determination of herniation nature and a performance of 62.8% for hernia direction determination. This issue arises from the inability of the models to adequately generalize the data used because of the insufficient data or a lack of correlation between the available data and desired outcomes.

4. Discussion

In the current era, ML is becoming increasingly prominent in our lives. ML is expected to enable software-powered robots to perform tasks that were once carried out by humans. Moreover, the utilization of ML in fields such as healthcare and education are steadily on the rise (1, 2). Thus, software programs capable of simultaneously comparing multiple parameters and producing results will replace humans who obtain experience and skills gained through formal education. This could make it possible for a general practitioner triaging in an emergency room or an individual overseeing healthcare insurance expenses to be replaced by such programs in the near future. Thus, we need to determine the usability of ML within the realm of healthcare and how it should be employed. Our study significantly contributes to literature by investigating the potential contribution of ML in assessing spinal pathologies and shedding light on how effective it can be in this context.

Cervical trauma is less commonly encountered compared to trauma at other spinal regions. Factors that could lead to spinal pathologies, such as obesity, pregnancy, and heavy lifting, have a lesser impact on the cervical region than on the other spinal regions (5). Due to its simpler dynamic function, the cervical spine is more amenable to investigation for disc degeneration and herniation using ML systems than the other spinal regions (3). The most influential factors in this regard include the age, dominant extremity, nature and side of the disc herniation, and level of disc herniation. In our study, we initially statistically compared these data in patients and assessed their comparability with those in literature. Subsequently, we compared these findings with the results obtained via ML.

Takahashi et al. reported a higher prevalence of CDH, especially at the C6-7 level, in individuals with left-hand dominance (6). Kang et al. reported that factors such as dominant extremity, age, disc level, and nature play a crucial role in disc herniation and emphasized the need for further research to support these findings. They attributed this to patients unknowingly placing their non-dominant arm higher while writing with their dominant hand and more frequently bearing loads on the dominant hand (7). However, in our study, we did not find a significant correlation between dominant extremity and disc herniation, both during statistical analyses and within the ML system.

As people age, an increase in spinal degeneration, calcification, and the presence of pathologies such as osteophytes often become inevitable (8). This standard knowledge for medical practitioners is easily predictable due to its frequent occurrence in educational and practical applications. In our statistical study, there was a statistically significant correlation between aging and the occurrence of calcified CDH. However, in ML, achieving a high level of accuracy, often close to 90%, is crucial. For instance, health insurance companies aim to streamline expenses by eliminating unnecessary tests for their customers. In such a scenario, a computer programmer without basic medical knowledge could monitor the expenditures of insured customers. Similarly, medical students lacking sufficient clinical experience could benefit from similar support programs.

Our AI-supported study results indicate that age and gender heavily impact disc nature (8). In the test models, a significant decrease did not occur when only age and gender parameters were used. Thus, our original intention, to establish a correlation between the arm pain direction and hand dominance and the direction of CDH, might not hold true based on the study findings. This insight offers valuable guidance for further exploration and analysis of spinal pathologies.

This study focused on CDH. AI-supported applications do not yet appear to be suitable for practical use in this field due to the lack of sufficient data and correlations as well as inadequate diversity in the data utilized. However, AI-supported applications hold promise for the future. The appropriate identification, collection, and processing of relevant data will pave the way for more accurate results. This study has identified important foundation points to be considered in future AI-supported studies on the detection of CDH nature and direction.

Furthermore, changes in the diversity and quantity of collected data can lead to different results in other classification models. Models developed using deep learning algorithms could yield entirely different outcomes for parameters that are considered independent.

This study serves as a valuable reference for forthcoming AI-supported research on the detection of CDH nature and direction. Adjustments in the diversity and quantity of collected data, as well as the application of deep learning algorithms, could potentially yield diverse and impactful outcomes.

Conflict of interest
The authors declared no conflict of interest.

Funding
No funding was used for the study.

Acknowledgments
None to declare.

Authors' contributions
Concept: E.Y., Design: E.Y., Data Collection or Processing: E.Y., Analysis or Interpretation: Y.A., Literature Search: Y.A., Writing: E.Y., Y.A.,

Ethical Statement
The study was approved by the Alanya Alaaddin Keykubat University Ethics Committee (No: 11/06; approval date: 14.06.2023)

References

1. McDermott MBA, Wang S, Marinsek N, Ranganath R, Foschini L, Ghassemi M. Reproducibility in machine learning for health research: Still a ways to go. Sci Transl Med. 2021; 13(586): eabb1655. doi: 10.1126/scitranslmed.abb1655.
2. Yin Z, Sulieman LM, Malin BA. A systematic literature review of machine learning in online personal health data. J Am Med Inform Assoc. 2019; 26(6): 561-576. doi: 10.1093/jamia/ocz009.
3. Karabacak M, Margetis K. Machine learning-based prediction of short-term adverse postoperative outcomes in cervical disc arthroplasty patients. World Neurosurg. 2023; 1878-8750(23): 00796-00799. doi: 10.1016/j.wneu.2023.06.025.
4. Vujović, Ž. Classification model evaluation metrics. Int J Adv Comp Sci Appl. 2021; 12(6): 599-606.
5. Scarcia L, Pileggi M, Camilli A, Romi A, Bartolo A, Giubbolini F, et al. Degenerative disc disease of the spine: from anatomy to pathophysiology and radiological appearance, with morphological and functional considerations. J Pers Med. 2022; 12(11): 1810. doi: 10.3390/jpm12111810.
6. Takahashi Y, Yasuhara T, Kumamoto S, Yoneda K, Tanoue T, Nakahara M, et al. Laterality of cervical disc herniation. Eur Spine J. 2013; 22(1): 178-182. doi: 10.1007/s00586-012-2565-8.
7. Kang KC, Lee HS, Lee JH. Cervical radiculopathy focus on characteristics and differential diagnosis. Asian Spine J. 2020; 14(6): 921-930. doi: 10.31616/asj.2020.0647.
8. Fine N, Lively S, Séguin CA, Perruccio AV, Kapoor M, Rampersaud R. Intervertebral disc degeneration and osteoarthritis: a common molecular disease spectrum. Nat Rev Rheumatol. 2023; 19(3): 136-152. doi: 10.1038/s41584-022-00888-z.