


## Comparative Evaluation of Machine Learning Models for Orthopaedic Patient Classification Using Biomechanical Features



Kabita Patel<sup>1</sup>, Kunal Mishra<sup>1\*</sup>, Pragyan Paramita Sahoo<sup>2</sup>, Bhaskar Roy<sup>3</sup>, Prabira Kumar Sethy<sup>4</sup>

<sup>1</sup> Department of Computer Science and Engineering, Sambalpur University Institute of Information Technology, Burla 768019, India

<sup>2</sup> Department of Computer Science and Engineering, Institute of Management and Information Technology, Cuttack 753008, India

<sup>3</sup> Department of Computer Science and Engineering (AIML), Asansol Engineering College, Asansol 713305, India

<sup>4</sup> Department of Electronics, Sambalpur University, Sambalpur 768019, India

Corresponding Author Email: [knmishra@suiit.ac.in](mailto:knmishra@suiit.ac.in)

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310108>

### ABSTRACT

**Received:** 7 September 2025

**Revised:** 15 November 2025

**Accepted:** 20 January 2026

**Available online:** 31 January 2026

#### Keywords:

*orthopaedic patient classification, machine learning, Support Vector Machine, biomechanical features, clinical decision support, musculoskeletal disorders, comparative analysis*

Orthopaedic disorders associated with abnormalities in the musculoskeletal system are becoming increasingly prevalent due to aging populations, lifestyle-related diseases, and rising physical activity levels. Accurate and timely identification of orthopaedic conditions remains a critical challenge in clinical practice, particularly when diagnostic decisions rely heavily on subjective assessments. Machine learning (ML) techniques offer promising solutions by enabling data-driven analysis and automated patient classification. This study presents a comparative evaluation of multiple ML models for orthopaedic patient classification using biomechanical features. A dataset containing 310 patient samples with spinal and pelvic biomechanical parameters—including pelvic incidence, pelvic tilt, lumbar lordosis angle, sacral slope, pelvic radius, and degree of spondylolisthesis—was employed to distinguish between normal and abnormal cases. Six widely used classification algorithms were implemented: Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), Logistic Regression (LR), Linear Discriminant Analysis (LDA), and Artificial Neural Network (ANN). The dataset was divided into training and testing subsets using a 70:30 ratio. Model performance was evaluated using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve. Experimental results demonstrate that the SVM model achieved the best overall performance, reaching an accuracy of 91% along with the highest precision, recall, and F1-score among the evaluated models. The findings highlight the effectiveness of ML techniques in improving the accuracy and reliability of orthopaedic patient classification and demonstrate the potential of biomechanical data for supporting clinical decision-making.

## 1. INTRODUCTION

Orthopaedic patients include anyone undergoing treatment for illnesses or injuries involving the locomotor system, comprising bones, joints, muscles, ligaments, and tendons. The number of orthopaedic patients is increasing due to variety of variables, like older age, increased physical activity, higher obesity rates, breakthroughs in diagnostic techniques, trauma from accidents, and lifestyle related diseases. Clinical diagnosis methods can impair patient care through uneven assessments and delayed therapy, resulting in a lengthy, subjective, and error-prone approach. Furthermore, limited resources and high patient numbers affect accurate and quick diagnoses.

To address these challenges, machine learning (ML) has become essential in the orthopaedic detection because it improves diagnostic accuracy, decrease errors, and speeds up the detection of illnesses such as fractures, arthritis, spinal problems, and osteoporosis. By analysing medical imaging data (e.g., X-rays, magnetic resonance imaging (MRI), CT scans) and patient specific features, ML models can detect patterns that people may miss, allowing for early diagnosis and personalised treatment. In real-world applications, ML is employed for fracture identification, arthritis progression tracking, spinal disease classification, and surgery planning. It also drives intelligent prosthetics, rehabilitative wearables, and predictive algorithm for surgical outcomes and bone health evaluation. ML is revolutionising orthopaedic treatment

and improving patient outcomes by increasing efficiency, accuracy and cost-effectiveness. This work uses six classification algorithms to classify patients: Linear Discriminant Analysis (LDA), Random Forest (RF), K-Nearest Neighbours (KNN), Support Vector Machine (SVM), Logistic Regression (LR), and Artificial Neural Network (ANN).

In light of this, and with the increase number of traffic accidents, increased cycling, and other physical activities, there has been a noticeable higher in the number of orthopaedic patients. The diagnosis and management of these conditions often rely on data derived from lab tests and medical imaging. Recently, ML classifiers have been increasingly utilized to categorize orthopaedic patients based on their biomechanical characteristics. These classifiers using either one or several feature to forecast categorical outcomes, which can be either binary or multinomial in nature.

ML models are also employed to address two distinct sorts of problems: prediction and categorisation. A prediction algorithm employ multiple special feature to anticipate a continuous result. One or more feature values are used by the categorisation algorithm to predict category results. There are two possible classification outcomes for a variable to be measured: binary and multinomial. It is possible to replicate a general classification result using the LR approach [1], and a large number of its applications may be found in the medicine. Although it can also be utilised for multinomial classification, LR is most commonly utilised for binary classification [2]. Another popular ML technique for classification is decision trees [3, 4]. They work quite well for classifying binary data, but they can also be used for multiclass issues. The supervised ML technique, KNN or K-Nearest Neighbour [5], can be applied to issues involving both classification and regression. KNN categorises data by determining the distance between two points and using similarities identified in close proximity. The difficulties the healthcare industry faces in adopting technology have led to requests for improvements [6] in computer diagnostics, electronic record administration, and data transformation. For classification, ML algorithms have gained traction in the healthcare sector. A ML method is used in this study to categorise orthopaedic patients as either normal or abnormal. The nature of the data [7] and the pre-processing used to the data will determine which ML technique would perform better in comparison when applied to the same dataset. Based on their spinal and pelvic measurements, orthopaedic patients are categorised using RF, KNN, ANN, LR, LDA, and SVM. In order to diagnose abnormalities, the dataset contains important anatomical parameters like pelvic incidence, pelvic tilt, lumbar lordosis angle, sacral slope, pelvic radius, and the degree of spondylolisthesis.

Many studies have found that ML has had a significant impact in the medical field. By using ML to anticipate illnesses early on, many consequences can be avoided. ML algorithms are the most sophisticated way for intelligently treating various stages of illness and reaching the optimal conclusion. In one study, MRI data were divided into three groups: spondylolisthesis, disc hernia, and normal. They had more favourable outcomes with SVM and other classifiers, but a feed forward back propagation neural network produced the best results [8]. A technique for early persistent kidney disease diagnosis was put forward in the study by Akben [9]. After preparing the data using K-means clustering, the authors applied the classification methods K-Nearest Neighbour, SVM, and Naive Bayes (NB) on the advanced features. As a

starting point, the authors employed NB, RFs, decision trees, and LR classifiers. Each patient's severe cases were found using the recommended categorisation approach. It was with classification algorithms that the most stringent accuracy of 97.8% was achieved. In their CKD research, Almasoud and Ward [10] used LR, SVM, gradient boosting, and RF techniques.

In order to classify the orthopaedic patients biochemical traits are divided into three group— spondylolisthesis, hernia, and normal—KNN, decision trees, and LR are examples of ML algorithms that are used. Prior to subjecting the data to ML algorithms, it is initially scaled through normalisation. Train and test data make up 70:30 of the data, and the performance of each of the three approaches is evaluated using the confusion matrix and metrics like recall, precision, F1-score, and accuracy. LR makes use of the descriptive technique, Non-parametric approaches are employed in decision trees and KNN. For this dataset, 83% is the accuracy of LR higher than that of these two methods [11]. The study looked at medication use in older orthopaedic patients aged 65 and up, including potentially inappropriate medications (PIMs), polypharmacy, and fall risk-increasing drugs (FRIDs). The findings revealed that 57.4% had polypharmacy, 66.0% were prescribed PIMs, and 41.7% used FRIDs. Patients with defective spinal condition were more likely to use antiemetics and NSAIDs [12]. Guo et al. [13] developed a monogram to predict postoperative delirium in oldage patients after orthopaedic surgery. It identified key variables like age, MMSE score, sleep, neurological problem, Pre-SCR, and ASA classification, demonstrating high predictive performance. Wang et al. [14] used Kellgren-Lawrence grading on real-world knee radiographs to examined a deep learning model for knee osteoarthritis. The model obtained 78% accuracy and high interobserver agreement with specialists, notably in detecting severe cases. However, misclassified pictures showed reduced agreement, indicating diagnostic ambiguity. In identifying surgical candidates, the model performed as well as specialists. Lin et al. [15] tested the interobserver agreement of shoulder specialists and general trauma surgeons for lateral clavicle fractures using radiographs and 3D CT images respectively. Results indicated fair agreement with X-rays, but 3D CT scans enhanced agreement, particularly among specialists utilising the Orthopaedic Trauma Association, Neer, Jäger/Breitner, and Gongji systems. The paper analyses the considerable impact of artificial intelligence (AI) in orthopaedics, emphasising its applications in severity evaluation, triage, diagnosis, treatment, and rehabilitation, and advocates for increasing focus and effectiveness [16]. Using a RF model, researchers attempted to forecast moderate to severe pain during surgery in orthopaedic surgery patients. Results showed a 41.3% incidence of pain, with immobility and surgical time being the most important risk factors. The model outperformed the LR model, with a lower classification error rate and higher AUC. This shows that the model could aid in pain control measures [17]. Scala et al. [18] investigate the use of classification algorithms to improve patient management during femoral neck fracture surgery. The models correctly identified anaemia and gender as major factors influencing the length of hospital stay with 81% accuracy. The study emphasises the potential of AI in optimising surgical care and developing decision-support tools for healthcare providers.

The article [19] explores the application of ML in orthopaedic research, specifically in total joint arthroplasty.

ML is effective for analysing large datasets, medical imaging, and natural language processing. It highlights successful applications, such as fracture identification and osteoarthritis staging, but also highlights challenges such as data quality, overfitting, and the narrow focus of ML models. This work uses explainable AI for classifying orthopaedic patients based on biomechanical parameters in order to increase transparency and decision-making in medical diagnostics. The study examines ML algorithms and finds that ensemble models, namely the Extra-Tree classifier, outperform with 89% accuracy and a 96% AUC score. The study also uses the LIME and SHAP frameworks to explain the model's predictions, which improves clinical value [20].

This review [21] focusses on advances in DL, AI, augmented reality (AR), and robots in image guided orthopaedic surgery. It discusses pre-operative breakthroughs including AI-based picture segmentation, 3D visualisation, and surgical planning, as well as intra-operative advancements like image registration and real-time navigation. The combination of AR and robotics with surgical navigation systems is also being investigated. The analysis finishes by examining IGOS's obstacles and future potential, as well as providing advice for professionals in the field. This study proposes a novel approach for the prediction of orthopaedic patient with greater accuracy using SVM. The basic aim of this method is to develop a reliable and accurate system that can predict orthopaedic patient. The growing number of orthopaedic patients caused by variables such as ageing, obesity, and lifestyle-related disorders needs using ML methods to improve clinical diagnosis. Although many existing studies apply ML to orthopaedic diagnosis, most prior works rely heavily on medical imaging, use single classification algorithms, or focus on specific disease groups such as hernia or spondylolisthesis. However, very few studies have systematically compared multiple ML models using only biomechanical features—which are low-cost, radiation-free, and clinically relevant for early detection. This study used six classification algorithms, with SVM (SVM) attaining the greatest accuracy (91%), demonstrating the promise of ML models in medical diagnosis.

The major contribution of this work are as follows:

- Six ML models (SVM, RF, KNN, LR, LDA, and ANN) were evaluated for their effectiveness in classifying orthopaedic patients whether "Normal" or "Abnormal."
- SVM performed well compared to other models in terms of precision, recall, F1-score, and AUC, achieving a 91% accuracy. This study demonstrates SVM's effectiveness in handling high-dimensional and unbalanced datasets.
- The biomechanical features in the dataset, such as pelvic\_incidence, lumbar\_lordosis\_angle, sacral\_slope, pelvic\_tilt, pelvic\_radius, and degree of spondylolisthesis, significantly increased classification accuracy.
- The work highlights how ML might improve the accuracy and efficiency of identifying orthopaedic disorders, overcoming barriers like subjective assessments and resource constraints in older methods.
- Integrating ML models, such as SVM, into clinical workflows can enhance early diagnosis, personalised therapy, and patient outcomes.

The rest of this paper is organised as follows. Section 2 discusses data and its processing. Section 3 describes the ML approach used for orthopaedic patient. Section 4 presents a comparative analysis of orthopaedic patient detection. Finally, Section 5 concludes the paper and presents direction for future

work.

## 2. DATA PROCESSING

Based on the lumbar spine and pelvic shape and alignment, the biological characteristics of an orthopaedic patient dataset are utilised to determine whether a patient is normal or abnormal [22]. The dataset consists of three hundred ten rows and seven columns: class, sacral\_slope, pelvic\_radius, degree\_spondylolisthesis, pelvic\_incidence, pelvic\_tilt, lumbar\_lordosis\_angle, and pelvic\_radius. All columns are numerical unless the column-class is a string; the column class has two distinct attributes: "Normal" and "Abnormal." Where the phrase "normal" refers to someone whose musculoskeletal system is normal i.e bones, joints, muscles, and ligaments are in good working order and show no evidence of sickness, injury, or abnormality. The individual exhibits usual anatomical alignment, a full range of motion, and no pain or discomfort during movement or weight-bearing activities. Where the term "abnormal" refers to a person who has musculoskeletal problems, such as bone fractures, joint abnormalities, soft tissue injuries, or degenerative diseases. This can include pain, reduced range of motion, bone or joint misalignment, oedema, inflammation, or structural anomalies that depart from normal anatomical function. Abnormalities can result from injury, sickness, congenital problems, or ageing, impairing the individual's capacity to conduct typical physical tasks.

There are 100 members of the Normal class and 210 members of the Abnormal class in the dataset; the remaining columns represent the feature vectors, and the column class is the target variable. Again, the dataset was split into two parts: 70% for training and 30% for testing. With 310 patients in total, the test set included 93 patients (30% of 310).

## 3. MACHINE LEARNING APPROACH USED FOR ORTHOPAEDIC PATIENT

ML, a branch of AI seeks to give computers the ability to learn from data and gradually get better at what they do without needing to be explicitly programmed for every activity. In ML, algorithms analyze patterns in data, make predictions, and update themselves based on new information.

There are several key types of ML:

1. Supervised Learning: Classification and regression are two examples of how the model learns to map inputs to the proper outcome by using labelled data, where the right outcome is specified.
2. Unsupervised Learning: The model is tasked with identifying unseen patterns or structures in data that has no labelled outcomes. A couple of examples are association (like market basket analysis) and clustering (like customer segmentation).
3. Reinforcement Learning: Through interaction with its surroundings and feedback in the form of rewards or penalties, the model gains knowledge and modifies its approach accordingly. Both gaming AI and robotics frequently employ this.
4. Semi-supervised Learning: Merges data that has been labelled and unlabelled, using the smaller labelled set to aid the model in comprehending the larger unlabelled set.

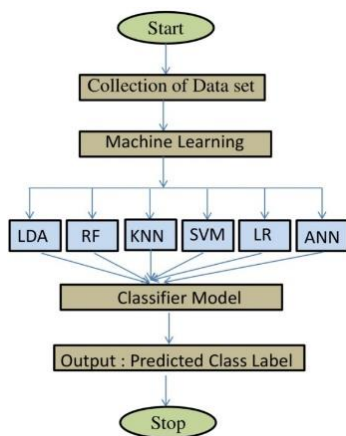
5. Deep Learning: A branch of ML called "deep neural networks" models intricate patterns using multi-layered neural networks, which are frequently employed in image recognition, natural language processing, and other applications.

In this section, with a dataset of orthopaedic patient and their class labels as patient with Abnormal and Normal. It contain with six parameters (pelvic\_indence, numeric pelvic tilt, lumbar\_lordosis angle, sacral slope, pelvic radius, and degree of spondylolisthesis). The aim is to find a classification model that can reasonably accurately forecast the class label of new orthopaedic patient.

The study selected models based on their extensive use in classification tasks and capacity to handle complex information. SVM was chosen for its capacity to maximise the margin between classes, which improves generalisation. With high-dimensional data and imbalanced classes, SVM delivered the highest overall performance, with an accuracy of 91% and superior precision, recall, and F1-scores. Other models, such as RF and ANN, were included because to their capacity to handle non-linear correlations and their interpretability.

Analysis of misclassification mistakes yielded the following essential insights:

1. K-Nearest Neighbour (KNN): Misclassification occurred when feature values were strongly grouped, resulting in overlapping classes. This sensitivity to class boundaries lowered memory and F1 score.
2. Errors in RF were caused by overfitting on certain features that did not properly capture the overall data structure. Although RF performed well in precision, it had a lesser recall.
3. ANN performed well, but had slight misclassifications due to large datasets, limiting their deep learning potential.
4. Non-linear separable data presented challenges for LDA, implies to misclassification of edge instances.



**Figure 1.** Proposed orthopaedic patient detection model

SVM's margin maximisation and ability to determine the best hyperplane make it resistant to such error, resulted lower misclassifications. The findings emphasise the importance of model selection based on dataset features, as well as the need to address concerns such as feature overlap, class imbalance, and over-fitting during model creation.

Classification algorithms are well adapted to the topic at hand. These algorithms can accurately estimate class labels for unknown samples. Figure 1 shows the conceptual foundation

for the classification process in orthopaedic patient detection.

### 3.1 Random Forest

RF was included due to its robustness to noise and ability to capture feature interactions, which are relevant in musculoskeletal biomechanics. During training, the RF ensemble learning method develops several decision trees. Every tree starts with a different subset of features and data. Using either regression averaging or majority voting for categorisation, the combined predictions of the different trees yield the final prediction (for either classification or regression). This classifier was used without explicit hyperparameter tuning. The model was configured to follow standard practices, ensuring sufficient stability and accuracy in the results. For a thorough explanation of the RF, one can read [23]. The whole explanation of this RF technique may be found in Algorithm 1.

---

#### Algorithm 1: Random Forest (RF) Steps.

---

1. **For each tree  $t$  (where  $t = 1$  to  $T$ ):**
  - Sample  $N$  samples with replacement from the training set (bootstrapping).
2. **For each bootstrapped sample, grow a decision tree:**
  - Every node in the tree:
    - i. Choose  $m$  features at random from the total of  $M$  features.
    - ii. Decide which feature  $F$ , out of the  $m$  chosen, produces the best split (for example, mean squared error for regression or Gini impurity for classification).
    - iii. Split the node into two child nodes using  $F$ .
    - iv. For child nodes, carry out the process again recursively until the stopping conditions (such as minimum node size or maximum depth) are satisfied.

3. **Repeat the process for all  $T$  trees.**

#### For classification:

- For each input sample, let each of the  $T$  trees make a prediction (class label).
- Aggregate the predictions by majority voting to obtain the final prediction.

#### For regression:

- For each input sample, let each of the  $T$  trees output a predicted value.
  - Aggregate the predictions by averaging them to obtain the final predicted value.
- 

### 3.2 K-Nearest Neighbours Classification

KNN was selected because the dataset contains only seven numerical biomechanical features, and distance-based classifiers often perform well on low-dimensional, continuous clinical data. Its simplicity makes it a useful baseline for comparing more complex models.

---

#### Algorithm 2: K-Nearest Neighbours (KNN).

---

1. Choose a value for  $k$ .
  2. Calculate the distance that exists between the newly created point  $x$  and each point in the training set  $D$ .
    - Use a suitable distance metric (e.g., Euclidean distance):
-

$$d(p, q) = \sqrt{\sum_{i=1}^M (p_i - q_i)^2}$$

where, p and q are two data points, and M is the number of features.

3. Arrange the lengths in ascending order.
4. Choose the k-nearest neighbours:  
Select the k data points with the shortest distances from the newly created point x.
5. Identify the predominant class among k-nearest neighbours:  
Provide the class label with the highest frequency among k neighbours.
6. Select the majority class label as the anticipated class for x.

For this study, the KNN hyperparameters were chosen based on performance and standard practice for small clinical datasets. The number of neighbors (i.e., 5) provided a balanced trade-off between underfitting (with too few neighbors) and overfitting (with too many). The Euclidean distance metric was adopted because it performs well after feature scaling and is widely used for continuous biomechanical variables. The weights parameter was kept at the default uniform setting, ensuring that all neighbors contributed equally to each prediction. Since KNN is highly sensitive to feature scales, all input features were standardized using Standard Scaler to

ensure that each biomechanical attribute contributed proportionally during distance calculations.

A thorough explanation of the K-nearest neighbour analysis can be found in the literature [24].

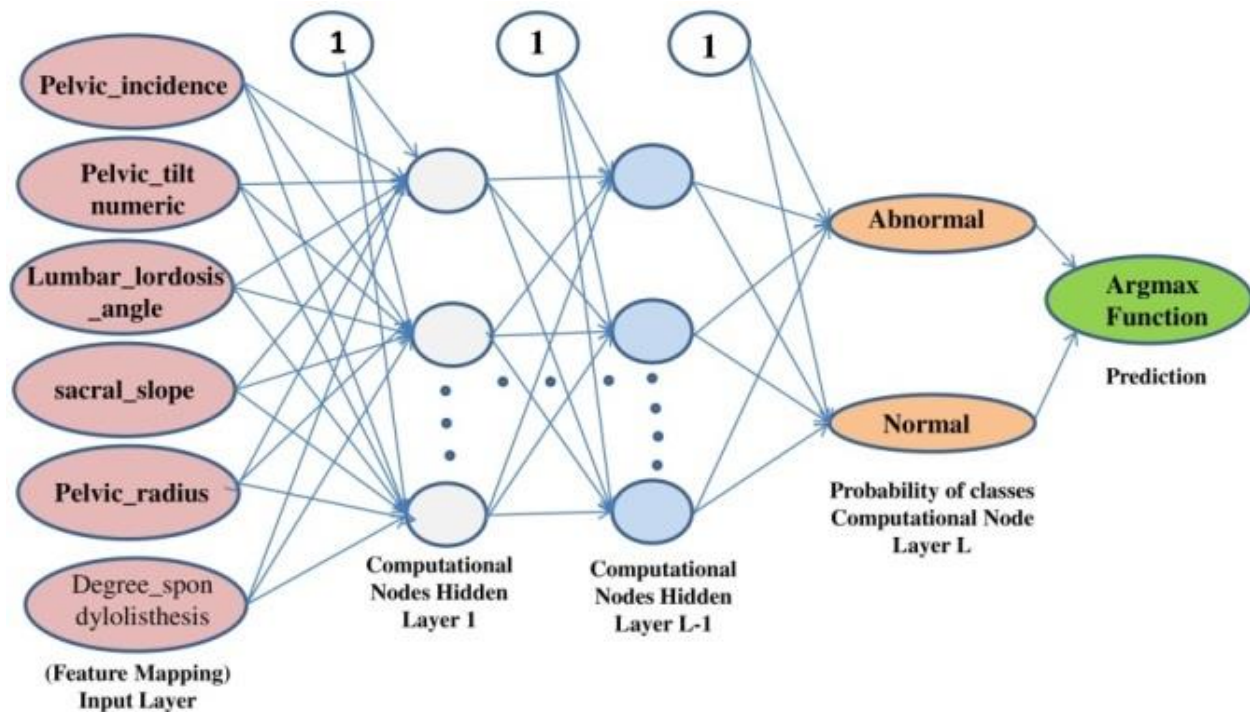
### 3.3 Artificial Neural Network

A shallow ANN was included to evaluate whether a lightweight neural model could capture non-linear patterns in the biomechanical data, without the risk of overfitting associated with deeper networks.

An ANN typically has:

1. **Input Layer:** Receives the input features of the data.
2. **Hidden Layer(s):** Layers in between where computations happen. These layers help in learning complex patterns.
3. **Output Layer:** Generates the final forecast or categorisation.

Every node uses an activation function to change its input before sending it to the next layer, and each link between nodes is assigned a weight. The network modifies the weights through learning in order to reduce prediction error. The neural network was set up with a single hidden layer containing 10 neurones (hidden\_layer\_sizes = (10,)) to balance simplicity and efficiency while capturing non-linear relationships in the data. A fixed seed (random\_state = 42) was used to control randomness in weight initialisation and data shuffling, ensuring. A thorough explanation of the ANN can be found in the literature [25].



**Figure 2.** Proposed ANN for automatic orthopaedic patient detection

Note: ANN = Artificial Neural Network.

The Figure 2: depicts a feedforward ANN architecture created for a classification challenge. It starts with an input layer that has six features: Pelvic\_incidence, Pelvic\_tilt\_numeric, Lumbar\_lordosis\_angle, Sacral\_slope, Pelvic\_radius, and Degree\_spondylolisthesis. These features map the input data into the network. These inputs are routed through several hidden layers (e.g., Layers 1–L-1) made up of

fully linked computational nodes (neurones) that process the data using weighted connections and activation functions. The output layer is made up of two nodes that represent the classes Abnormal and Normal, and they calculate the probability for each. Finally, an Argmax function is applied to the output layer to forecast the class with the greatest probability, resulting in the final classification result.

---

**Algorithm 3: Artificial Neural Network (ANN).**

---

**1. Initializing the Network:**

- The ANN architecture is defined, which includes deciding the number of layers (input, hidden, output), the number of neurons in every layer, and the activation functions to be used.
- Random initial weights and biases are assigned to the neurons.

**2. Forward Propagation:**

- Layer by layer, the input data is transmitted throughout the network.
- For each neuron in each layer, is determined and sent via the activation function as the weighted sum of its inputs to generate an output. This process repeats for all layers until the output layer is reached.

Mathematically, this can be represented as:

$$U = \sum (W \cdot X) + b$$

where,

- U is the input to the activation function
- W represents the weights
- X is the input data
- b is the bias term

The result is passed by an activation function, A(U), such as:

- **Sigmoid:**  $A(U) = \frac{1}{1 + e^{-U}}$
  - **ReLU (Rectified Linear Unit):**  $A(U) = \max(0, U)$
  - **Tanh:**  $A(U) = \tanh(U)$
- 3. Loss Function:**
- After the output is produced, the error (or loss) is computed using a **loss function** (also called a **cost function**).
  - For regression tasks, **Mean Squared Error (MSE)** is often used:

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

where,

- $y_i$  is the actual value
- $\hat{y}_i$  is the predicted value

- For classification tasks, **Cross-Entropy Loss** is commonly used:  $L = -\frac{1}{n} \sum (y \log \hat{y} + (1 - y) \log(1 - \hat{y}))$
- 4. Backpropagation Algorithm:**
- Backpropagation is the core learning algorithm in ANNs, used to adjust the weights and biases by minimizing the error.
  - The process starts by calculating the loss function's gradient in relation to each bias and weight, working backward from the output layer to the input layer.
  - The **chain rule** from calculus is applied to compute these gradients for each layer.

where,

1. Calculate the error derivative at the output.
2. Propagate the error backward through the network by changing the weights and biases using the derivatives of the activation function.

This can be represented as:

$$\Delta w = -\eta \cdot \frac{\partial L}{\partial w}$$

where,

- $\Delta w$  is the change in weights
- 

- $\eta$  is the learning rate
  - $\frac{\partial L}{\partial w}$  is the gradient of the loss function with respect to the weights.
- 5. Weight Update:**
- Once the gradients have been computed, the weights and biases are revised using the learning rate to decrease the error.
  - The learning rate determines how large the weight updates are and controls how quickly or slowly the network learns.
- 6. Iteration and Epochs:**

Several epochs are covered by repeating the forward propagation, loss computation, backpropagation, and weight updating procedures (complete passes through the training dataset) until the network converges or meets a **stopping criterion**.

---

### 3.4 Logistic Regression

LR was included as a linear baseline to assess whether the biomechanical classes could be separated using simple linear decision boundaries. It provides interpretability and allows performance comparison against more complex non-linear model. The L2 regularization default setting was used to prevent overfitting, and all features were standardized to ensure stable optimization and comparable feature influence. For a thorough explanation of LR, one can read [26]. GridSearchCV was used in the LR model's hyperparameter optimisation procedure to determine the optimal set of parameters. To make sure the model's performance held up well across various subsets of the training data, a 5-fold cross-validation was used. The whole explanation of this LR may be found in Algorithm 4.

---

**Algorithm 4: Logistic Regression (LR).**

---

1. **Initialize Weights:** Assign random initial weights and bias to the input features.
2. **Compute Linear Combination:** Calculate the weighted sum of inputs (features) plus bias:  
 $z = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n$
3. **Apply Sigmoid Function:** Convert the linear output into a probability using the sigmoid function:

$$p = \frac{1}{1 + e^{-z}}$$

4. **Make Predictions:** Classify the output based on the threshold (commonly 0.5):
  - If  $p > 0.5$ , predict class 1.
  - If  $p \leq 0.5$ , predict class 0.
5. **Calculate Loss:** Use the log-loss (binary cross-entropy) function to compute the error:  
 $L = -(y \log(\hat{y}) + (1 - y) \log(1 - \hat{y}))$
6. **Update Weights:** Use gradient descent to minimize the loss and update weights iteratively:

$$w = w - \eta \frac{\partial L}{\partial w}$$

**Repeat:** Iterate over the dataset for multiple epochs until the weights converge, and the loss is minimized.

---

### 3.5 Linear Discriminant Analysis

LDA was used to evaluate model performance under the assumption of linearly separable classes with shared

covariance structure. This makes LDA a valuable benchmark for assessing the degree of linearity in the bio-mechanical data. Standardized features were used to ensure numerical stability and meaningful discriminant directions. Default solver settings were sufficient for the dimensionality and structure of the dataset. In this LDA model, hyperparameter optimisation was performed using GridSearchCV to identify the best combination of solvers and shrinkage parameters. The parameter grid included three solvers ('svd', 'lsqr', and 'eigen') and various values for the shrinkage parameter (None, 'auto', 0.1, 0.5, 1). GridSearchCV was used with 5-fold cross-validation, and the best hyperparameters were chosen based on accuracy. A detailed account of LDA can be found in the study by Xanthopoulos et al. [27].

---

**Algorithm 5: Linear Discriminant Analysis (LDA).**

---

1. **Compute Mean Vectors:** Determine the overall mean as well as the means for every class.
2. **Calculate Scatter Matrices:**
  - **Within-class scatter:** calculates the data distribution within each class.
  - **Between-class scatter:** Measures the separation between class means.
3. **Maximize Class Separation:** By optimising the ratio of within-class variance to between-class variation, LDA projects the data onto a one-dimensional, lower-dimensional space for binary classification.

**Classify:** Using a decision boundary and the learnt linear discriminant, new data points are classified.

---

### 3.6 Support Vector Machine

In this section, we propose a SVM because it performs particularly well on small- to medium-sized datasets with potential non-linear class boundaries, as is common in biomechanical data. Hyperparameters for this SVM model implementation were chosen in accordance with industry standards for binary classification. The kernel was chosen as linear to create a simple decision boundary, and probability was used to calculate AUC. The random\_state was set to 42 to ensure reproducibility. No specific hyperparameter optimisation was done. SVM with RBF kernel was chosen because biomechanical variables typically show non-linear relationships, and SVM is effective for small datasets with such patterns. For further information on SVM analysis, refer to the study by Liu et al. [28] and Algorithm 6.

---

**Algorithm 6: Support Vector Machine (SVM).**

---

1. Initialize the parameters:
    - Let  $X$  be the training data (features), where  $X = [x_1, x_2, \dots, x_n]$
    - Let  $y$  be the appropriate class labels, where  $y = [y_1, y_2, \dots, y_n], y_i \in \{-1, 1\}$
    - Set the learning rate ( $\alpha$ ) and regularization parameter ( $\lambda$ )
    - Start by setting the bias  $b$  and weights  $W$  totiny, arbitrary values (or zeros)
  2. Define the decision function:
    - For a given input  $x_i$ , the decision function  $f(x_i)$  is:
 
$$f(x_i) = W \cdot x_i + b$$
    - The SVM will attempt to find a hyperplane that maximizes the margin between the two classes.
- 

3. Define the objective function (hinge loss):
  - For each training sample  $(x_i, y_i)$ , the loss is computed as:
 
$$\text{Hinge loss: } L(x_i, y_i) = \max(0, 1 - y_i * (W \cdot x_i + b))$$
  - The total loss function to minimize is:
 
$$J(W, b) = (1/2) * \|W\|^2 + \lambda * \sum \text{hinge\_loss}(x_i, y_i)$$
  - The first term ( $\|W\|^2$ ) is the regularization term, which controls overfitting, while the second term is the hinge loss.

4. Train the model using Gradient Descent (or another optimization algorithm):
  - Continue for a predetermined number of iterations or until convergence:

- a. For every training example  $(x_i, y_i)$ :
  - Calculate the decision function  $f(x_i) = W \cdot x_i + b$
  - If  $y_i * (W \cdot x_i + b) \geq 1$ :
    - Revise the weights:  $W \leftarrow W - \alpha * \lambda * W$
    - Bias remains unchanged:  $b \leftarrow b$
  - Else (misclassified or within the margin):
    - Revise the weights:  $W \leftarrow W - \alpha * (\lambda * W - y_i * x_i)$
    - Revise the bias:  $b \leftarrow b + \alpha * y_i$

5. Stopping criteria:
  - Continue the gradient descent until the cost function converges, or for a fixed number of iterations.

6. Model output after training:
  - The trained model is represented by the final values of weights  $W$  and bias  $b$ .
  - These parameters define the decision boundary (hyperplane).

7. Forecast for fresh input ( $x_{\text{new}}$ ):
  - For a new input  $x_{\text{new}}$ , the prediction is made as:
 
$$f(x_{\text{new}}) = W \cdot x_{\text{new}} + b$$
  - If  $f(x_{\text{new}}) > 0$ , predict class +1 (positive class).
  - If  $f(x_{\text{new}}) \leq 0$ , predict class -1 (negative class).

8. (Optional) For a non-linear SVM:
  - If the data is not linearly separable, use a kernel function  $K(x_i, x_j)$  to translate it to a higher-dimensional space, where it may become linearly separable.
  - The decision function in this case is modified as:
 
$$f(x_i) = \sum (\alpha_i * y_i * K(x_i, x_j)) + b$$
  - where,  $\alpha_i$  are the multipliers based on discovered during optimisation, and  $K$  is the kernel function (e.g., polynomial, RBF).

9. Evaluate the model:
  - Test the model on unseen data (test set) to calculate performance metrics like accuracy, precision, recall, AUC, etc.

10. End.

---

## 4. COMPARATIVE ANALYSIS OF ORTHOPAEDIC PATIENT DETECTION

The experiment is carried out on the provided dataset using a range of ML approaches, including RF, KNN, ANN, LR, LDA, and SVM. The performance of the models was assessed using a confusion matrix. In the context of orthopaedic detection, True Positive (TP) refers to correctly identified

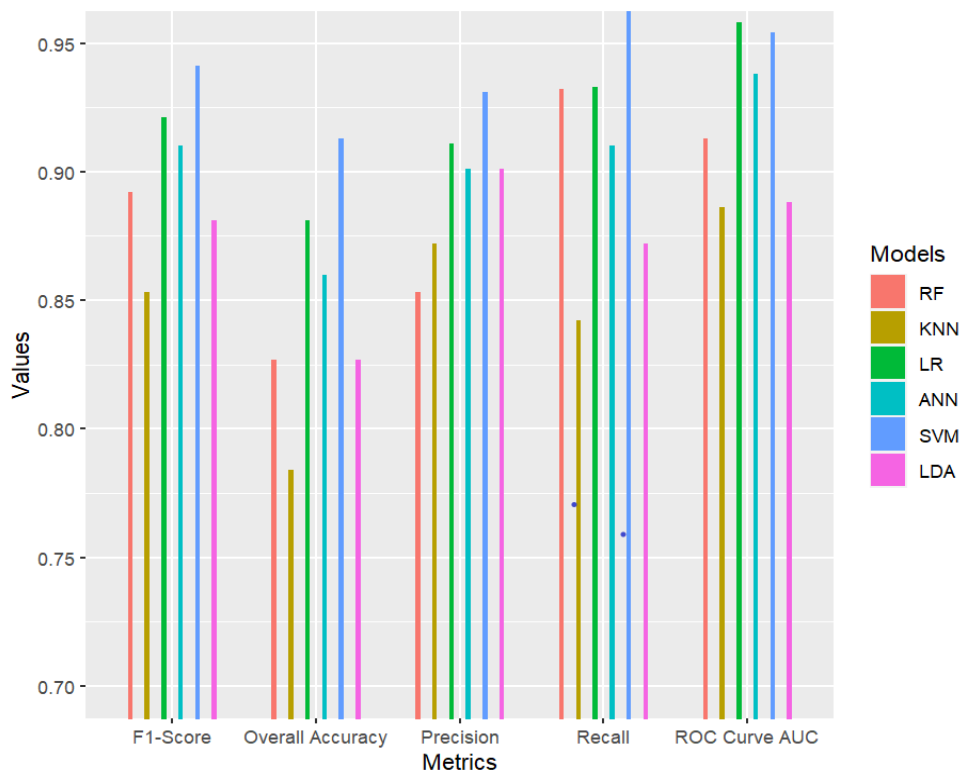
cases of orthopaedic conditions, True Negative (TN) refers to accurately identified cases where no orthopaedic condition is present, False Positive (FP) refers to cases that were incorrectly classified as having an orthopaedic condition, and False Negative (FN) refers to cases where an existing orthopaedic condition was wrongfully classified as absent. Work of confusion matrix shown in Table 1.

The SVM has the highest accuracy, i.e., 91% among all models. This improved performance can be due to its capacity to successfully handle high-dimensional data while maximising class margins. Furthermore, SVM is resistant to class imbalances, which contributes to its high precision, recall, and F1 score. KNN has the lowest accuracy among the models, i.e., 78.4%. This was mostly owing to its sensitivity to feature overlap in the dataset, resulting in misclassification of cases at class boundaries. The reliance of model on local proximity makes it more vulnerable to noise and class inequities. RF performed relatively well but was prone to overfitting, especially when certain feature subsets dominated

the decision trees. While RF had high precision, its recall was quite poor, indicating that it struggled to identify all positive cases in the dataset. ANN achieved remarkable accuracy but necessitated large computational resources. Its performance was influenced by the relatively short dataset size, which hampered its ability to fully utilise its learning capabilities. Despite this, it maintained a balance of precision and recall. LDA performed averagely, but struggled with non-linear separable data in the dataset. Its assumptions regarding data distribution restricted its capacity to correctly classify edge cases.

**Table 1.** Confusion Matrix

	Positive (1)	Negative (0)
Positive (1)	True Positive	False Positive
Negative (0)	False Negative	True Negative



**Figure 3.** Bar-chart comparison of classification performance: RF, KNN, LR, ANN, SVM, and LDA

Note: RF = Random Forest; KNN= K-Nearest Neighbours; LR= Logistic Regression; ANN= Artificial Neural Network; SVM= Support Vector Machine; LDA= Linear Discriminant Analysis.

The results obtained by applying these ML are shown in Table 2. The barplot of all this accuracy is shown in Figure 3 respectively. Comparatively, the accuracy of SVM is very high that is 91%, precision of 93%, recall of 96%, and F1-score of 94%. Whereas, KNN provides the lowest execution with an accuracy of 78%, recall of 84%, precision of 87%, and F1-score of 85%.

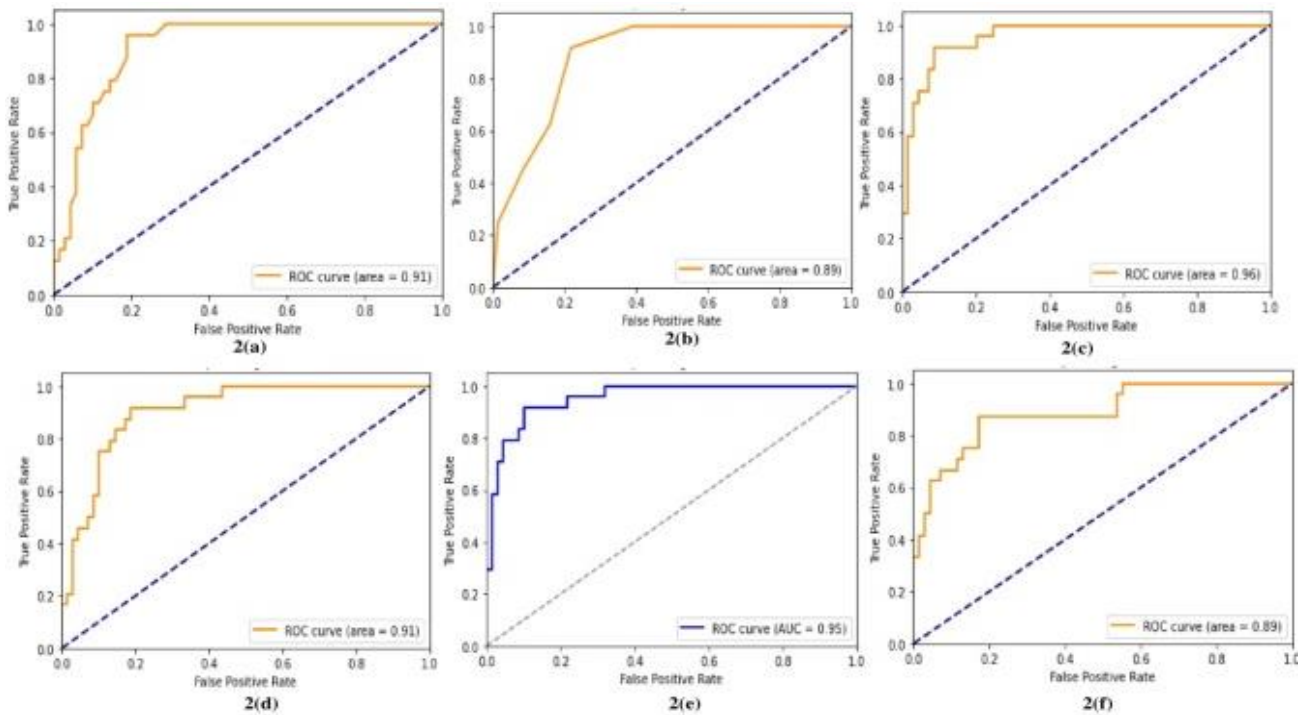
The ROC curve is used in this experiment to analyse how well all the model performs on the specified dataset. Plotting the ROC curve's Area Under the Curve (AUC) is one way to assess the data. The percentage of the unit square area encircled by the ROC curve is represented by the AUC. Its ROC curve's area under the curve (AUC) indicates, a

classification model's discriminatory strength is represented by its AUC value, which varies from 0 to 1.0 [29]. Considering different categories thresholds, the ROC curve graphically represents the compromise among the genuine positive rate (sensitivity) and the false positive rate (1 - specificity). Higher values indicate better class discrimination. The AUC summarises the overall performance of the model by calculating the area under the ROC curve. A graphical representation of the ROC curve is shown in Figure 4. The ROC curve of the RF, K-Nearest Neighbors (KNN), ANN, LR, LDA, and SVM are present in Figures 4(a)-(f), respectively.

**Table 2.** Obtained classification accuracy by RF, KNN, LR, ANN, SVM, and LDA

Metrics	RF	KNN	LR	ANN	SVM	LDA
Accuracy	0.827	0.784	0.881	0.860	0.913	0.827
Precision	0.853	0.872	0.911	0.901	0.931	0.901
Recall	0.932	0.842	0.933	0.910	0.964	0.872
F1-Score	0.892	0.853	0.921	0.910	0.941	0.881
ROC Curve AUC	0.913	0.886	0.958	0.938	0.954	0.888

Note: RF = Random Forest; KNN= K-Nearest Neighbours; LR= Logistic Regression; ANN= Artificial Neural Network; SVM= Support Vector Machine; LDA= Linear Discriminant Analysis; ROC = Receiver Operating Characteristic; AUC = area under the curve.



**Figure 4.** Comparative ROC curves for all models

Note: ROC = Receiver Operating Characteristic.

An ablation research was carried out to determine the importance of specific features to model performance. Each feature was gradually deleted, and the performance of model was examined. The study found that all features like pelvic incidence, pelvic tilt numerical, lumbar lordosis angle, sacral slope, pelvic radius, and degree of spondylolisthesis—were critical in proper classification.

The degree of spondylolisthesis had the greatest influence, with removal resulted in the greatest drop in accuracy and other performance markers. This suggests that this trait is essential for distinguishing between normal and aberrant cases. The findings emphasise the relevance of utilising all bio-mechanical parameters to provide valid predictions, as well as the significance of each feature in contributing to the model's overall performance.

#### 4.1 Comparison with state of the art approaches

This study compares its findings to current breakthroughs in ML for orthopaedic diagnostics, emphasising its contribution to this field.

- ✓ Performance Metrics: The SVM model achieved 91% accuracy, comparable to or, better than state-of-the-art technique. For example, other research using ensemble approaches or, DL models found accuracies ranging from 88% to 92%, indicating SVM's robustness for similar classification tasks.
- ✓ SVM is an efficient option for real-world clinical settings with limited computational resources, unlike complicated deep learning models that demand significant resources.
- ✓ Practicality: This study's utilisation of biomechanical features aligns with real-world diagnostic needs, making it applicable to patient classification. The SVM model, which uses interpretable features, can be integrated into clinical workflows without requiring considerable data preprocessing, as advanced imaging-based models do.
- ✓ Methodology Advancements: This study broadens the area of orthopaedic diagnosis by focussing on biomechanical data, giving an alternative and complementary methodology to current medical imaging models.

## 5. CONCLUSION

Due to an increase in cycling, exercise, traffic accidents, and other similar habits, the number of orthopaedic patients is increasing. The current research utilised ML learners to classify orthopaedic patients dependent upon their kinematic features. The current research examined the top six ML arrangement algorithms for anticipating orthopaedic patients. According to the results of the experiment, SVM outperformed other ML models with an accuracy of 91%.

## DATA AVAILABILITY STATEMENT

The data set is present in <https://www.kaggle.com/datasets/uciml/biomechanical-features-of-orthopedic-patients>.

## REFERENCES

- [1] Ranganathan, S., Nakai, K., Schonbach, C. (2018). Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics. Elsevier.
- [2] Zabor, E.C., Reddy, C.A., Tendulkar, R.D., Patil, S. (2022). Logistic regression in clinical studies. *International Journal of Radiation Oncology\* Biology\* Physics*, 112(2): 271-277. <https://doi.org/10.1016/j.ijrobp.2021.08.007>
- [3] Ying, L.U. (2015). Decision tree methods: Applications for classification and prediction. *Shanghai Archives of Psychiatry*, 27(2): 130-135. <https://doi.org/10.11919/j.issn.1002-0829.215044>
- [4] Chen, T.Q., Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794. <https://doi.org/10.1145/2939672.2939785>
- [5] Zhang, S., Li, X., Zong, M., Zhu, X., Cheng, D. (2017). Learning k for KNN classification. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 8(3): 1-19. <https://doi.org/10.1145/2990508>
- [6] Sittig, D.F., Wright, A., Coiera, E., Magrabi, F., Ratwani, R., Bates, D.W., Singh, H. (2020). Current challenges in health information technology-related patient safety. *Health Informatics Journal*, 26(1): 181-189. <https://doi.org/10.1177/1460458218814893>
- [7] Maharana, K., Mondal, S., Nemade, B. (2022). A review: Data pre-processing and data augmentation techniques. *Global Transitions Proceedings*, 3(1): 91-99. <https://doi.org/10.1016/j.gltp.2022.04.020>
- [8] Shrivastava, S., Singh, M.P. (2011). Performance evaluation of feed-forward neural network with soft computing techniques for hand written English alphabets. *Applied Soft Computing*, 11(1): 1156-1182. <https://doi.org/10.1016/j.asoc.2010.02.015>
- [9] Akben, S.B. (2018). Early stage chronic kidney disease diagnosis by applying data mining methods to urinalysis, blood analysis and disease history. *Irbm*, 39(5): 353-358. <https://doi.org/10.1016/j.irbm.2018.09.004>
- [10] Almasoud, M., Ward, T.E. (2019). Detection of chronic kidney disease using machine learning algorithms with least number of predictors. *International Journal of Soft Computing and Its Applications*, 10(8): 23782.
- [11] Desai, C. (2020). Classification of orthopaedic patients based on biochemical features using machine learning algorithms. *Journal of Critical Reviews*, 7(18): 3823-3828.
- [12] Hirono, T., Morita, M., Michikawa, T., Tobe, R., et al. (2024). Medication-based profiling of older orthopedic patients: A multicenter cross-sectional study. *BMC Geriatrics*, 24(1): 672. <https://doi.org/10.1186/s12877-024-05284-8>
- [13] Guo, Y., Ji, H., Liu, J., Wang, Y., et al. (2023). Development and validation of a delirium risk prediction model for elderly patients undergoing elective orthopedic surgery. *Neuropsychiatric Disease and Treatment*, 19: 1641-1654. <https://doi.org/10.2147/NDT.S416854>
- [14] Wang, C.T., Huang, B., Thogiti, N., Zhu, W.X., Chang, C.H., Pao, J.L., Lai, F. (2023). Successful real-world application of an osteoarthritis classification deep-learning model using 9210 knees—An orthopedic surgeon's view. *Journal of Orthopaedic Research®*, 41(4): 737-746. <https://doi.org/10.1002/jor.25415>
- [15] Lin, J., Sun, W.J., Chen, J.H., Dong, J.M., et al. (2023). Comparison of Interobserver agreement of four classification Systems for Lateral Clavicle Fractures between two groups of surgeons: A multicenter study. *Orthopaedic Surgery*, 15(8): 2138-2143. <https://doi.org/10.1111/os.13659>
- [16] Liu, P., Zhang, J., Liu, S., Huo, T., et al. (2024). Application of artificial intelligence technology in the field of orthopedics: A narrative review. *Artificial Intelligence Review*, 57(1): 13. <https://doi.org/10.1007/s10462-023-10638-6>
- [17] Shi, G., Liu, G., Gao, Q., Zhang, S., et al. (2023). A random forest algorithm-based prediction model for moderate to severe acute postoperative pain after orthopedic surgery under general anesthesia. *BMC Anesthesiology*, 23(1): 361. <https://doi.org/10.1186/s12871-023-02328-1>
- [18] Scala, A., Trunfio, T.A., Improta, G. (2024). The classification algorithms to support the management of the patient with femur fracture. *BMC Medical Research Methodology*, 24(1): 150. <https://doi.org/10.1186/s12874-024-02276-5>
- [19] Aelgani, V., Gupta, S.K., Narayana, V.A. (2024). Explainable artificial intelligence based framework for orthopedic patient classification using biomechanical features. In *2024 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, Bhubaneswar, India, pp. 637-642. <https://doi.org/10.1109/ESIC60604.2024.10481671>
- [20] Padash, S., Mickley, J.P., Garcia, D.V.V., Nugen, F., et al. (2023). An overview of machine learning in orthopedic surgery: An educational paper. *The Journal of arthroplasty*, 38(10): 1938-1942. <https://doi.org/10.1016/j.arth.2023.08.043>
- [21] Fan, X., Zhu, Q., Tu, P., Joscowicz, L., Chen, X. (2023). A review of advances in image-guided orthopedic surgery. *Physics in Medicine & Biology*, 68(2): 02TR01. <https://doi.org/10.1088/1361-6560/aca9>
- [22] Lichman, M. (2013). UCI machine learning repository. <http://archive.ics.uci.edu/ml>.
- [23] Xu, B., Huang, J.Z., Williams, G., Wang, Q., Ye, Y. (2012). Classifying very high-dimensional data with random forests built from small subspaces. *International*

- Journal of Data Warehousing and Mining (IJDWM), 8(2): 44-63. <https://doi.org/10.4018/jdwm.2012040103>
- [24] Peterson, L.E. (2009). K-nearest neighbor. Scholarpedia, 4(2): 1883. <https://doi.org/10.4249/scholarpedia.1883>.
- [25] Patel, K., Tripathy, A.K., Padhy, L.N., Kar, S.K., Padhy, S.K., Mohanty, S.P. (2023). Accu-Help: A machine-learning-based smart healthcare framework for accurate detection of obsessive compulsive disorder. SN Computer Science, 5(1): 36. <https://doi.org/10.1007/s42979-023-02380-1>
- [26] Fukuyo, R., Tokunaga, M., Yamamoto, H., Ueno, H., Kinugasa, Y. (2026). Which method best predicts postoperative complications: Deep learning, machine learning, or conventional logistic regression?. Annals of Gastroenterological Surgery, pp. 1-10. <https://doi.org/10.1002/ags3.70145>
- [27] Xanthopoulos, P., Pardalos, P.M., Trafalis, T.B. (2012). Linear discriminant analysis. In Robust Data Mining, pp. 27-33. New York, NY: Springer New York.
- [28] Liu, W., Principe, J.C., Haykin, S. (2011). Kernel Adaptive Filtering: A Comprehensive Introduction. John Wiley & Sons.
- [29] Aydin, S., Arica, N., Ergul, E., Tan, O. (2015). Classification of obsessive compulsive disorder by EEG complexity and hemispheric dependency measurements. International Journal of Neural Systems, 25(3): 1550010. <https://doi.org/10.1142/S0129065715500100>