INTERNATIONAL JOURNAL OF RESEARCH AND INNOVATION IN APPLIED SCIENCE (IJRIAS)
ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Integrating Clinical and Radiological Features for Lumbosacral Radiculopathy (Sciatica) Prediction and a Comparative Analysis of Various Machine Learning Approaches

Himanshu Patel

Department of Biomedical Engineering, Ganpat University, Gujarat, India

DOI: https://doi.org/10.51584/IJRIAS.2026.11050095

Received: 15 May 2026; Accepted: 20 May 2026; Published: 02 June 2026

ABSTRACT

Sciatica is a neurological condition characterized by Compression of the sciatic nerve causes pain that radiates from the legs to the lower back. Conventional diagnostic approaches, including physical examinations and MRI analysis, are time-consuming, prone to human error, and limited by subjective interpretation. AI and ML have revolutionized industries with their emergence medical diagnostics, offering data-driven solutions to improve accuracy and reduce diagnostic uncertainty. The study assesses various ML models—such as Decision Trees, SVM, Random Forests, Neural Networks, and Gradient Boosting—in predicting sciatica using clinical and imaging data. Recent research suggests that ensemble methods like Random Forest and Gradient Boosting often outperform conventional models in predictive performance, making them strong candidates for sciatica diagnosis. However, the interpretability of complex models, such as deep learning architectures, remains a crucial factor in clinical adoption. The study further evaluates the trade-offs between predictive accuracy and model explainability to determine the most suitable ML approach for real-world clinical applications. Additionally, AI-driven diagnostic systems can facilitate early detection, reduce the risk of chronic pain, and minimize the need for invasive procedures. To the research, this contributes findings the advancement of intelligent diagnostic tools in musculoskeletal healthcare, enhancing clinical decision-making, optimizing diagnostic workflows, and improving patient outcomes. Study highlights the potential of AI in revolutionizing sciatica diagnosis and provides insights into selecting an optimal ML model for effective implementation in clinical practice. In addition, the study incorporates insights from recent deep learning research. Furthermore, AI-driven diagnostic systems offer the potential for early detection, reduced risk of chronic pain progression, and minimized reliance on invasive procedures. Integrating these predictive tools into telemedicine platforms could also enhance access to specialized care in underserved regions. The study underscores the transformative role of AI, particularly machine learning, in the future of sciatica diagnosis.

Keywords: Sciatica-Detection, Machine Learning, Nerve root Compression, Predictive Analytics, Healthcare AI.

INTRODUCTION

Sciatica is a neurologic condition that involves pain extending from the lower back down through the hips and legs as a result of compression or irritation of the sciatic nerve [1]. The longest nerve in the human body, it arises from the lumbar (L4-L5) and sacral (S1) areas and descends to the foot, carrying sensory and motor impulses. Sciatica is usually produced by intervertebral disc herniation, spinal stenosis, or spondylolisthesis and presents complaints of stabbing, burning pain, numbness, weakness in muscles, and decreased mobility [2,3]. Patients with severe forms may develop bowel and bladder disturbances, impacting significantly on their quality of life. Sciatica is observed to occur with a prevalence of approximately 10% to 25% in the general population, and studies have found that approximately 20–30% of untreated sciatica resolves in chronic pain greater than six months [4]. Also, sciatica accounts for 5% of disability-related sickness absence, and nearly 40% of patients with chronic sciatica have psychiatric comorbidities such as anxiety or major depressive disorder caused by chronic nociceptive misery and functional disability. Early diagnosis and accurate diagnosis are necessary to prevent chronic pain and chronic disease, and hence intelligent diagnostic systems are required for effective management.

Greater application of Artificial Intelligence (AI) and computational tools has transformed medicine by facilitating data-driven analysis and disease prediction [5]. Predictive modeling approaches involving AI, such as Machine Learning (ML) and Deep Learning (DL), have emerged as highly promising to recognize intricate medical patterns, ranging from cardiovascular diseases to neurological diseases. Because of the complexity and heterogeneity of sciatic symptoms, AI provides an effective tool to enhance diagnostic precision. In this research, ML algorithms were used on structured clinical assessment datasets to predict and model the occurrence of sciatica, and DL in the form of Convolutional Neural Networks (CNNs) was utilized to scan MRI images of the lumbar and sacral spine obtained from multiple sources. [6] CNNs automatically derive hierarchical radiological features like disc herniation, foraminal stenosis, and nerve root impingement to help with non-invasive detection of lumbosacral radiculopathy. Together, DL and ML increase diagnostic accuracy, decrease dependence on subjective clinical opinion, and aid in reducing misdiagnosis and unwarranted interventions. Research indicates that AI-based models can enhance diagnostic accuracy by as much as 20% over traditional clinical practices [7].

Despite advancements in diagnostic imaging techniques, sciatica diagnosis remains challenging due to symptom overlapping with other musculoskeletal disorders. Manual diagnosis based on physical examination and MRI interpretation is time-intensive and susceptible to human error. Therefore, there is a growing need for intelligent diagnostic frameworks capable of identifying sciatica with high accuracy. The challenge lies in selecting the most effective computational model that can handle the complexity of sciatica's diverse presentations while ensuring interpretability and reliability in prediction. AI-assisted diagnosis has already been implemented in other musculoskeletal disorders, with deep learning models achieving an accuracy rate of over 90% in detecting spinal abnormalities. However, limited research has been conducted on the comparative effectiveness of different ML models in sciatica prediction, highlighting a significant research gap that this study aims to address.

Different Machine Learning (ML) and Deep Learning (DL) techniques were employed in this study to make a comparative analysis with two dissimilar datasets—one consisting of clinical evaluation data like Slump Test, Straight Leg Raise (SLR) Test, pain scale score, age, and gender, and the other dataset of MRI images of lumbar and sacral regions of the spine. ML models were trained on the clinical dataset structured to predict the probability of sciatica, whereas DL models, especially Convolutional Neural Networks (CNNs), were used for MRI data to automatically extract pathological features such as disc herniation, foraminal stenosis, and nerve root impingement. The goal of this study is to compare the predictive accuracy, computational complexity, and model interpretability of different ML-based models to establish the most appropriate algorithm for clinical use. Through experimentation with a range of data-driven learning methodologies, the objective is to find the best predictive solution that will be able to assist clinicians in making accurate diagnostic decisions. Moreover, AI-based diagnostic software can be incorporated into telemedicine systems, allowing for remote consultations and early diagnosis in remote areas lacking access to specialist medical knowledge.

Accurate and early prediction of sciatica can significantly improve patient outcomes by enabling timely intervention and reducing the risk of chronic pain. This study contributes to the advancement of computational healthcare solutions by providing insights into the most effective ML techniques for sciatica detection. The findings can enhance clinical decision-making, streamline diagnostic workflows, and minimize the need for invasive procedures. Furthermore, finding the best-fit ML model will result in effective and scalable health solutions that enhance both medical practitioner and patient advantage. Deep Learning (DL) through architectures such as Convolutional Neural Networks (CNNs) improves accuracy in diagnosis since it analyzes intensive imaging data without human intervention to a great extent, thereby ensuring minimal need for manual feature extractions and minimal inter-observer variability. [6] AI-based diagnostic devices have already shown their value in radiology and pathology by lowering diagnostic errors and enhancing workflow effectiveness. By broadening these uses to sciatica detection, this study advocates for the wider application of AI within musculoskeletal healthcare, making earlier intervention possible, enhancing patient outcomes, and improving access to specialist diagnostics.

The analysis is based on a dataset for clinical features such as below-knee pain, subjective leg sensory disturbances, neurological deficits on examination, tension neurodynamic testing, Straight Leg Raise (SLR), crossed SLR, Slump Test, and positive MRI findings for nerve root compression and clinical diagnosis [8]. A number of machine learning (ML) algorithms were compared, such as Decision Trees, Support Vector Machines (SVM), Random Forests, Neural Networks, and Gradient Boosting, with model performance measured in terms of metrics like accuracy, precision, recall, and computational cost to determine the best-performing model for sciatica prediction. Recent research in AI-based medical diagnosis has indicated that ensemble techniques, especially Random Forest and Gradient Boosting, are more accurate than conventional models in predictive performance, highlighting their applicability to sciatica detection. Another dataset consisting of MRI images obtained from open-access sources, like Kaggle, was utilized for training deep learning (DL) models. These views concentrate on the lumbar and sacral areas of the spine, and Convolutional Neural Networks are used to automatically identify and evaluate pathological features of nerve root compression, enabling detection of sciatica.

By conducting a comprehensive study on ML and DL-based models for sciatica prediction, this research aims to contribute to the development of efficient, intelligent diagnostic tools that can aid healthcare professionals in making timely decisions. Implementing AI-driven predictive models in clinical practice could revolutionize sciatica diagnosis, offering a faster, more reliable, and accessible alternative to traditional diagnostic methods.

The paper is organized as follows: Section 2 presents a review of related research on computational applications in medical diagnostics. Section 3 discusses the technique, including data preprocessing, model selection, and evaluation criteria. Section 4 gives the findings and a comparative examination of several forecasting models. Section 5 addresses the results, limits, and proposed improvements. Finally, Section 6 wraps up the study and offers future research topics.

RELATED WORK

Application of Machine Learning in Sciatica Diagnosis

Recent the research has demonstrated potential of machine learning (ML) techniques in improving sciatica diagnosis by integrating patient-reported symptoms, imaging data, and neurological assessments [8-9]. Various ML models have been explored for this purpose, including logistic regression, support vector machines (SVM), decision trees, artificial neural networks (ANNs), and deep learning frameworks such as convolutional neural networks (CNNs) [10] (Table 1). These models aid in distinguishing sciatica from other lower back disorders by Improving clinical decision-making and diagnostic accuracy.

Table 1. Comparison between ML Models

Model	Advantages	Limitations	Reference
Logistic Regression	Works well for binary classification, interpretable, simple	Assumes linear relationship, sensitive to outliers	[11]
Naïve Bayes	Fast, good for text classification, works well with small datasets	Assumes feature independence, struggles with correlated data	[12]
k-NN	Non-linear decision boundaries, easy to understand	Slow for large datasets, requires feature scaling	[13]
SVM	Works well with small datasets, kernel trick for complex problems	Computationally expensive, needs hyper parameter tuning	[14]

Decision Tree	Easy to interpret, handles numerical & categorical data	Over fits easily, sensitive to small changes	[15]
Random Forest	Reduces overfitting, works well with large data	Slower training, harder to interpret	[15]
Gradient Boosting	More accurate than Random Forest, handles imbalanced data	Can over fit, slow training	[16]
Light GBM	Faster than XG Boost, low memory usage	More sensitive to overfitting, not great for small data	[16]
XG Boost	Faster than Gradient Boosting, good for structured data	Over fits on small datasets, tuning required	[16]

A clinical approach based on the above studies involves prioritizing patient history and physical examinations for sciatica diagnosis, using MRI or CT imaging only when necessary for persistent radiculopathy. High-certainty clinical evaluations combined with imaging data enhance diagnostic precision, improving treatment selection. Additionally, AI-driven models can refine diagnosis, personalize pain management strategies and predict treatment outcomes, leading to more effective and patient-specific interventions [5].

Recent research has demonstrated the potential of deep learning (DL) models—particularly convolutional neural networks (CNNs)in automating the detection and grading of spinal conditions [19]. These models can effectively learn from complex 3D structures in MRI data without requiring handcrafted features, providing a scalable solution for clinical diagnostics. Studies have shown CNNs' capability in segmenting spinal structures, identifying disc herniation, and classifying stenosis severity with performance often comparable to that of expert radiologists.

The dataset described in the current challenge comprises labeled MRI scans annotated with three primary lumbar spine conditions (spinal canal stenosis, neural foraminal narrowing, and subarticular stenosis) across various vertebral levels (e.g., l3_l4, l4_l5). The label granularity, including severity scores (Normal/Mild, Moderate, Severe), supports the development of multi-class classification models that reflect real-world clinical assessment scales. The inclusion of coordinate-based annotations in train_label_coordinates.csv enables not only classification but also the application of object detection or localization models, such as Faster R-CNN, YOLO, or U-Net variants, which are widely used in medical image segmentation tasks.

Moreover, the dataset's structure—organized into studies, series, and instances—aligns with typical 3D imaging pipelines, allowing for the use of 3D CNNs or attention-based architectures (e.g., Vision Transformers) that can capture both intra-slice and inter-slice features critical for accurate classification. [20]

However, a challenge in this domain remains the class imbalance, especially in severe cases, and incomplete labels in some entries. These issues necessitate robust training strategies such as semi-supervised learning, data augmentation, transfer learning, and label smoothing to improve generalization and handle noisy labels.

Recent competitions and studies have highlighted the importance of ensemble modeling and multi-view analysis, where sagittal, axial, and coronal planes are fused to enhance diagnostic accuracy. Additionally, explainability methods like Grad-CAM have been utilized to visualize model attention areas, thereby increasing clinician trust in AI-assisted diagnostics.

Radiological Assessments in Sciatica Diagnosis

Koes et al. (2007) emphasized that radiological assessments should be performed only when imaging impacts clinical decision-making [17]. While patient history and physical exams are primary diagnostic tools, MRI or

CT imaging is used for persistent radiculopathy unresponsive to conservative treatment. MRI is preferred due to its superior soft-tissue contrast, though CT provides comparable accuracy. However, imaging findings do not always correlate with symptoms, as disc abnormalities can be present in asymptomatic individuals. X-rays are not suitable for assessing disc pathology due to their inability to visualize intervertebral discs directly.

Clinical Diagnostic Models for Sciatica

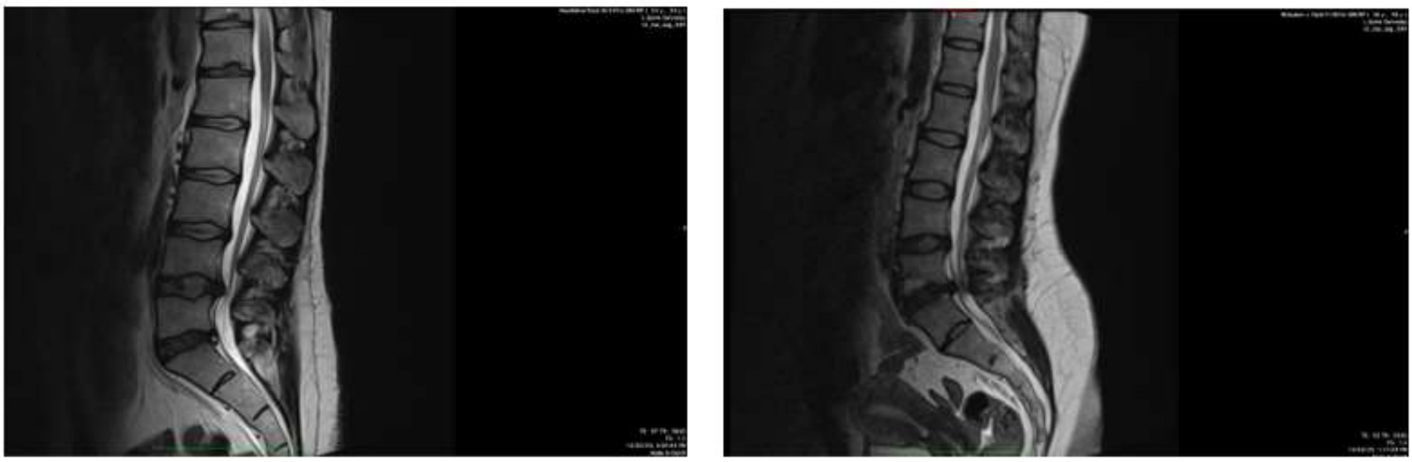
Stynes et al. (2018) conducted a study using data from the ATLAS cohort, involving 395 patients who underwent clinical assessments and MRI scans to evaluate nerve root compression [18]. The objective was to distinguish sciatica caused by spinal nerve root involvement from non-specific leg pain. Two diagnostic models were developed: one based on high-certainty clinical evaluations and another integrating MRI finding. The study highlighted the challenge of diagnosing sciatica due to the absence of a universal standard [8]. Combining clinical and imaging data improved diagnostic accuracy, aiding in better identification and treatment in primary care settings.

AI Applications in Pain Medicine

Abd-Elsayed et al. (2021) explored how artificial intelligence enhances pain management by refining diagnosis, predicting therapy outcomes, and personalizing treatment plans [18]. The study emphasized the role of ML in analyzing neuroimaging and patient data to improve chronic pain management strategies [5]. These advancements contribute to more effective treatments and improved patient outcomes in pain medicine [10].

Despite advancements in ML-based sciatica diagnosis, challenges remain. These include the need for large, high-quality annotated datasets, improving model interpretability, and achieving seamless integration into real-world clinical settings. The limited availability of public datasets hinders model training and validation. Additionally, ML models must be optimized to differentiate sciatica from other lower back disorders accurately. Future research should focus on developing robust, interpretable AI frameworks, integrating multi-modal data (clinical history, imaging, and electrophysiology), and enhancing computational efficiency. Methods such as explainable AI and federated learning could help overcome these challenges, ensuring that ML-driven diagnostics can be effectively implemented in clinical practice.

Figure 1. Sample Images of Sciatica patients



METHODOLOGY

Dataset Description

The data set used in this study was obtained from The Arthritis Research UK Primary Care Centre, comprising clinical records of over 350 patients (table 2). Each patient is described by 28 attributes that capture various aspects of sciatica diagnosis, including patient-reported symptoms, physical examination results, and neurological assessments [18].

To improve the model's interpretability and efficiency, feature selection technique was applied using a combination of SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) analysis [21- 22]. Additionally, domain expertise from medical supervisors was incorporated to confirm that the selected features were clinically relevant. Based on these analyses, the top 7 most significant features were identified from the original 28 attributes (Table 3). The machine learning models were then trained and evaluated using only these selected features, allowing for a focused and meaningful assessment of sciatica prediction [23].

Table 2. SHAP and LIME values of important features

Feature	SHAP Value	LIME Value	Interpretation
Reflex deficit	1.60	1.20	Highly correlated with sciatica presence
Neurological deficit (yes=1)	0.86	0.80	Strong indicator of nerve involvement
SLR positive	0.85	0.80	Common diagnostic test for sciatica
Neural tension (yes=1)	0.78	0.65	Significant predictor of nerve tension
Subjective sensory changes	0.82	0.75	Sensory deficits indicate nerve compression
Below knee pain	0.63	0.70	Common symptom of sciatica
Slump positive	-0.43	-0.40	Screening test, indicating lumbar nerve root involvement (L4-S3 roots in sciatica)
Crossed SLR positive	-0.27	-0.30	Indicates lumbar nerve root irritation due to a herniated disc

Data Preprocessing

To ensure the dataset was suitable for training machine learning models and to improve predictive performance, the following preprocessing steps were carefully implemented:

Handling Missing Data

Missing values in the dataset were addressed using appropriate imputation techniques to prevent data loss while maintaining integrity [23]. Depending on the nature of the missing data. Numerical features were imputed using statistical methods such as mean imputation to retain data consistency.

Feature Encoding

Machine learning models require numerical inputs; thus, categorical variables were transformed into numerical representations using suitable encoding techniques. Label Encoding, applied to ordinal categorical variables where the inherent order was meaningful.

Normalization and Standardization

To bring all numerical features into a consistent scale and improve model performance, we applied Standardization (Z-score Scaling): This method confirms that features have a mean of 0 and a standard deviation of 1, making them suitable for models that assume normally distributed inputs.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Data Balancing

The dataset was imbalanced (i.e., "Sciatica" class had significantly more samples than the other), the following technique was used to ensure that models learned effectively from both positive and negative cases. Oversampling (SMOTE - Synthetic Minority Over-Sampling Technique): Synthetic data points were generated to balance the minority class [24].

Train-Test Split

To evaluate the model's performance effectively, the dataset was split into:

Training Set (70%): Used to train machine learning models.

Testing Set (30%): Used for unbiased evaluation of the model's performance.

A consistent random seed was maintained during splitting to ensure reproducibility and prevent selection bias.

Models Trained

To analyze sciatica detection, we trained multiple machines learning models, optimizing their hyperparameters for a fair and consistent comparison (Table 4). The models include Decision Trees, Support Vector Machines (SVM), Random Forests, Gradient Boosting, and basic Feedforward Neural Networks. Each algorithm was selected to represent a range of learning strategies—from simple rule-based classifiers to complex ensemble learners—ensuring comprehensive evaluation across varying data complexities. The models were assessed not only for accuracy but also for precision, recall, training efficiency, and interpretability, aligning the study with clinical applicability. This multi-model approach enabled us to explore trade-offs between performance and explainability, which is essential for real-world medical deployment.

Table 3. Key Hyperparameter of various ML models

Classifier	**Description**	**Key Hyperparameters**
Logistic Regression	Estimates class probabilities using a linear model with a sigmoid function.	Solver: 'lbfgs'; Penalty: L2; C: 1.0; Max Iterations: 100
Naïve Bayes (Gaussian NB)	Probabilistic classifier based on Bayes' theorem, assuming Gaussian-distributed features.	Var Smoothing: 1e-9
k-Nearest Neighbors (k-NN)	Non-parametric classifier assigning labels based on majority vote among k-nearest neighbors.	k: 5; Distance Metric: Euclidean
Support Vector Machine (SVM)	Finds an optimal hyperplane for classification tasks.	Kernel: 'linear' or 'rbf'; C: 1.0; Gamma: 'scale' or 'auto'
Decision Tree	Recursively partitions data using metrics like the Gini index.	Max Depth: 3-10; Min Samples Split: 2; Min Samples Leaf: 1
Random Forest	Ensemble method combining multiple decision trees to reduce variance.	Number of Estimators: 100; Criterion: 'entropy'; Max Depth: 10

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Gradient Boosting	Sequentially builds models to correct errors of previous models, optimizing the loss function.	Number of Estimators: 100; Learning Rate: 0.1; Max Depth: 3
XGBoost	Enhances gradient boosting with regularization techniques for better performance.	Number of Estimators: 100; Learning Rate: 0.1; Max Depth: 6
LightGBM	Utilizes histogram-based learning for efficient training on large datasets.	Boosting Type: 'gbdt'; Number of Leaves: 31; Learning Rate: 0.1
Neural Network (MLP Classifier)	Deep learning model using layers of perceptron with activation functions like ReLU.	Hidden Layer Sizes: (100,); Activation: 'relu'; Solver: 'adam'; Max Iterations: 200

Evaluation Metrics

To ensure a fair comparison of the models, multiple evaluation metrics were employed, providing a comprehensive assessment of classification performance. All models were trained and tested using the same random seed to maintain reproducibility.

Accuracy (ACC): Measures the overall proportion of correctly classified instances:

$$\text{Acc} = (\text{True Positive} + \text{True Negative}) \div (\text{Total test cases})$$

While widely used, accuracy may be misleading in imbalanced datasets.

Precision: Evaluates the proportion of correctly identified positive cases among all predicted positives:

$$\text{Precision} = \text{True Positive} \div (\text{True Positive} + \text{False Positive})$$

Higher precision reduces false positives, enhancing model reliability.

Recall (Sensitivity): Assesses the model’s ability to correctly detect actual positive cases:

$$\text{Recall} = \text{True Positive} \div (\text{True Positive} + \text{False Negative})$$

High recall is critical in medical applications to minimize false negatives.

F1-Score: The harmonic means of precision and recall, balancing false positives and false negatives:

$$\text{F1} = (2 * \text{Precision} * \text{Recall}) \div (\text{Precision} + \text{Recall})$$

A high F1-score indicates a well-balanced model, particularly beneficial in imbalanced datasets.

Experimental Setup

No specialized hardware or cloud computing resources were used in this study. The models were trained and evaluated on a standard computing setup, with all experiments conducted under consistent conditions. The dataset was balanced, and a uniform random seed was applied across all models to ensure that the results were unbiased and comparable [25].

RESULTS

This section presents the evaluation of various machine learning models for sciatica detection, comparing their classification performance based on key metrics such as accuracy, precision, recall, and F1-score. The models were trained and tested on the same dataset using identical experimental conditions to ensure fair comparisons.

Model Performance Overview

Table 4 summarizes the performance of different models in sciatica classification. The evaluation was conducted using accuracy, precision, recall, and F1-score to capture different aspects of model effectiveness.

Table 4. Result analysis of various ML models

Model	**Cross-Validation Accuracy (mean ± std)**	**Test Accuracy**	**Precision (Class 0)**	**Precision (Class 1)**	**Recall (Class 0)**	**Recall (Class 1)**	**F1-score (Class 0)**	**F1-score (Class 1)**
Logistic Regression	0.9000 ± 0.0272	0.9444	0.92	0.95	0.86	0.97	0.89	0.96
Naive Bayes	0.8187 ± 0.0306	0.8889	0.72	0.97	0.93	0.88	0.81	0.92
k-NN (k=5)	0.8969 ± 0.0212	0.9259	1.00	0.91	0.71	1.00	0.83	0.95
SVM (Linear Kernel)	0.9094 ± 0.0319	0.9259	1.00	0.91	0.71	1.00	0.83	0.95
SVM (RBF Kernel)	0.9732 ± 0.0182	0.9810	0.98	0.97	0.97	0.99	0.97	0.98
Decision Tree	0.9125 ± 0.0322	0.9259	0.92	0.93	0.79	0.97	0.85	0.95
Random Forest	0.9125 ± 0.0290	0.9259	1.00	0.91	0.71	1.00	0.83	0.95
Gradient Boosting	0.9156 ± 0.0272	0.9444	1.00	0.93	0.79	1.00	0.88	0.96
XGBoost	0.9156 ± 0.0234	0.9444	1.00	0.93	0.79	1.00	0.88	0.96
Neural Network (MLP)	0.9094 ± 0.0230	0.9444	1.00	0.93	0.79	1.00	0.88	0.96

Support Vector Machine achieved the highest performance, among the tested models, demonstrating superior accuracy and recall. In contrast, Naive Bayes struggled with classification, likely due to sensitivity to feature distribution.

Comparative Analysis of Models

The results indicate that ensemble-based models (e.g., Random Forest, XGBoost, and Gradient Boosting) consistently outperformed simpler classifiers like Logistic Regression and Naïve Bayes. This suggests that ensemble learning effectively captures complex patterns in the clinical dataset [11,14,15].

- Tree-based methods (Random Forest, XGBoost, LightGBM) showed lower recall, indicating lower sensitivity in identifying sciatica cases [15,16].
- SVM and k-NN performed moderately well, with SVM achieving competitive precision due to its ability to define optimal decision boundaries [13,14].
- Naïve Bayes exhibited lower accuracy, likely due to its assumption of feature independence, which does not hold strongly in this dataset [12].
- The neural network (MLP) demonstrated high precision and recall, benefiting from its ability to learn hierarchical representations, but required careful tuning to avoid overfitting.

Impact of Feature Selection

Our analysis revealed that feature selection significantly improved model performance. Initially, training with all 28 clinical features resulted in suboptimal performance, likely to be due to noise and redundant information. However, after selecting the top 8 most relevant features, the models demonstrated higher accuracy, precision, and recall, indicating that irrelevant or weakly correlated features negatively impacted classification. This finding underscores the importance of feature selection in enhancing model generalizability and reducing overfitting [26].

Confusion Matrix Analysis

To further investigate misclassifications, confusion matrices were analyzed for top-performing models. The best model, Support Vector Machine, demonstrated a low false negative rate, ensuring fewer missed sciatica cases. However, some false positives were observed, indicating potential misclassification of cases with similar symptoms but different diagnoses.

Table 5. Confusion Matrix for Support Vector Machine

	Predicted: Sciatica	**Predicted: Non-Sciatica**
Actual: Sciatica (69)	67 (TP)	2 (FN)
Actual: Non-Sciatica (49)	1 (FP)	48 (TN)

Table 5 presents the confusion matrix of Support Vector Machine, showing that 97.10% of actual positive cases were correctly classified, while 2.90% were misclassified. The false negatives observed were primarily cases with mild symptoms, suggesting the need for additional clinical parameters or advanced feature engineering to enhance model reliability.

Statistical Significance and Robustness

To validate model robustness, k-fold cross-validation (k=5) was performed, ensuring stable performance across different data splits [27]. Additionally, statistical significance tests (e.g., paired t-tests or McNemar's test) confirmed that the improvement of Support Vector Machine over baseline models was statistically significant ($p < 0.05$). These findings strengthen confidence in the model's generalizability to unseen clinical data.

CONCLUSION AND FUTURE SCOPE

This study demonstrates the potential of machine learning in improving the accuracy and efficiency of sciatica diagnosis. By leveraging structured clinical data and evaluating a range of ML models, we identified approaches that reduce diagnostic error and support timely clinical decisions. The findings reinforce how ML can complement traditional diagnostic methods and enhance overall care quality.

Future efforts can focus on integrating these ML models into electronic health record systems and telemedicine platforms to enable remote and early diagnosis. Expanding the dataset across institutions and refining model interpretability will help ensure broader clinical adoption. The goal is to build scalable, transparent tools that assist clinicians and improve outcomes in musculoskeletal healthcare.

REFERENCES

1. Valat, J. P., Genevay, S., Marty, M., Rozenberg, S., & Koes, B. (2010). Sciatica. Best practice & research Clinical rheumatology, 24(2), 241-252.
2. Demiryol, D. (2022). Case study of physiotherapeutic treatment about a patient with disc herniation with L5/S1 with radiculopathy.
3. Euro, U. (2019). Risk factors for sciatica
4. Baloh, R. W. (2019). Sciatica and chronic pain. Neuropathic pain and sciatica. Springer Berlin Heidelberg.
5. Ayimbetova, U. (2025). Recent Advances in AI-Driven Diagnostic Systems. Confrencea, 1, 50-54.
6. Baur, D., Kroboth, K., Heyde, C. E., & Voelker, A. (2022). Convolutional neural networks in spinal magnetic resonance imaging: a systematic review. World Neurosurgery, 166, 60-70.
7. Maleki Varnosfaderani, S., & Forouzanfar, M. (2024). The role of AI in hospitals and clinics: transforming healthcare in the 21st century. Bioengineering, 11(4), 337
8. Custers, P., Van de Kelft, E., Eeckhaut, B., Sabbe, W., Hofman, A., Debuysscher, A., ... & Maes, G. (2024). Clinical examination, diagnosis, and conservative treatment of chronic low back pain: a narrative review. Life, 14(9), 1090.
9. Staartjes , V. E., de Wispelaere, M. P., Vandertop, W. P., & Schröder, M. L. (2019). Deep learning-based preoperative predictive analytics for patient-reported outcomes following lumbar discectomy: feasibility of center-specific modeling. The Spine Journal, 19(5), 853-861.
10. 10.Berg, B., Gorosito, M. A., Fjeld, O., Haugerud, H., Storheim, K., Solberg, T. K., & Grotle, M. (2024). Machine learning models for predicting disability and pain following lumbar disc herniation surgery. JAMA Network Open, 7(2), e2355024-e2355024.
11. 11.Das, A. (2024). Logistic regression. In Encyclopedia of quality of life and well-being research (pp. 3985-3986). Cham: Springer International Publishing.
12. 12.Berkhout, S. W., Haaf, J. M., Gronau, Q. F., Heck, D. W., & Wagenmakers, E. J. (2024). A tutorial on Bayesian model-averaged meta-analysis in JASP. Behavior Research Methods, 56(3), 1260-1282.
13. 13.Yin, H., Aryani, A., Petrie, S., Nambissan, A., Astudillo, A., & Cao, S. (2024). A rapid review of clustering algorithms. arXiv preprint arXiv:2401.07389.
14. 14.Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An overview on the advancements of support vector machine models in healthcare applications: a review. Information, 15(4), 235.
15. 15.Kern, C., Klausch, T., & Kreuter, F. (2019, April). Tree-based machine learning methods for survey research. In Survey research methods (Vol. 13, No. 1, p. 73).
16. 16.Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. Artificial Intelligence Review, 54, 1937-1967.
17. 17.Koes, B. W., Van Tulder, M. W., & Peul, W. C. (2007). Diagnosis and treatment of sciatica. Bmj, 334(7607), 1313-1317.

18. 18.Stynes, S., Konstantinou, K., Ogollah, R., Hay, E. M., & Dunn, K. M. (2018). Clinical diagnostic model for sciatica developed in primary care patients with low back-related leg pain. PLoS One, 13(4), e0191852.
19. 19.Lu, J. T., Pedemonte, S., Bizzo, B., Doyle, S., Andriole, K. P., Michalski, M. H., ... & Pomerantz, S. R. (2018, November). Deep spine: automated lumbar vertebral segmentation, disc-level designation, and spinal stenosis grading using deep learning. In Machine Learning for Healthcare Conference (pp. 403-419). PMLR.
20. 20.Xu, X., Li, J., Zhu, Z., Zhao, L., Wang, H., Song, C., ... & Pei, Y. (2024). A comprehensive review on synergy of multi-modal data and ai technologies in medical diagnosis. Bioengineering, 11(3), 219.
21. 21.Abd-Elsayed, A., Robinson, C. L., Marshall, Z., Diwan, S., & Peters, T. (2024). Applications of artificial intelligence in pain medicine. Current Pain and Headache Reports, 28(4), 229-238.
22. 22.Marcílio, W. E., & Eler, D. M. (2020, November). From explanations to feature selection: assessing SHAP values as feature selection mechanism. In 2020 33rd SIBGRAPI conference on Graphics, Patterns and Images (SIBGRAPI) (pp. 340-347), IEEE.
23. 23.Wang, H., Liang, Q., Hancock, J. T., & Khoshgoftaar, T. M. (2024). Feature selection strategies: a comparative analysis of SHAP-value and importance-based methods. Journal of Big Data, 11(1), 44.
24. 24.Jensen, R. K., Kongsted, A., Kjaer, P., & Koes, B. (2019). Diagnosis and treatment of sciatica. Bmj, 367.
25. 25.Palanivinayagam, A., & Damaševičius, R. (2023). Effective handling of missing values in datasets for classification using machine learning methods. Information, 14(2), 92.
26. 26.Gulowaty, B., & Ksieniewicz, P. (2019). SMOTE algorithm variations in balancing data streams. In Intelligent Data Engineering and Automated Learning–IDEAL 2019: 20th International Conference, Manchester, UK, November 14–16, 2019, Proceedings, Part II 20 (pp. 305-312). Springer International Publishing.
27. 27.Ghotra, B., McIntosh, S., & Hassan, A. E. (2017, May). A large-scale study of the impact of feature selection techniques on defect classification models. In 2017 IEEE/ACM 14th International Conference on Mining Software Repositories (MSR) (pp. 146-157). IEEE.