

"Artificial Intelligence In Degenerative Lumbar Spine Management: Machine Learning Models For Data-Driven Clinical Decision-Making"

Sanjay Venkatesha², Dr. Nithya M², Dr. Komal Prasad Chandrachari¹, Dr. Madhusudhan Bharatwaj N¹, Dr. Mohammed Faisal, Sumukh Prasad, Aditya Prasad³

¹Department of Neurosurgery, Narayana Institute of Neurosciences, Mazumdar Shaw Medical Center, Narayana Health, Bengaluru, Karnataka

²Department of Management Studies, National Institute of Technology, Calicut, Kerala, India

³Birla Institute of Technology and Science, Pilani – Hyderabad, Telangana

Corresponding Author's Email:

Sanjay Venkatesha- sanjay_p240179ms@nitc.ac.in

Dr Nithya M- nithyam@nitc.ac.in

Dr. Komal Prasad Chandrachari-komal.prasad.dr@narayanahealth.org

Dr. Madhusudhan bharatwaj.N - msbbilla@gmail.com

Dr Mohammed Faisal-faisal31621@gmail.com

Sumukh Prasad- sumukhprasad.email@icloud.com

Aditya Prasad - adityacademics@gmail.com

Address for correspondence: Dr. Komal Prasad Chandrachari, Clinical Lead, Department of Neurosurgery, Narayana Institute of Neurosciences, Mazumdar Shaw Medical Center, Narayana Health, Bengaluru 560099, Karnataka komal.prasad.dr@narayanahealth.org

Abstract

Background:

Artificial Intelligence (AI) is transforming medicine by enhancing diagnostic accuracy, treatment planning, and patient monitoring through advanced data analysis^{[1][2][3]}.

We conducted a study at Narayana Health, Bangalore, to evaluate the utility of Machine Learning (ML) models—specifically Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN)—in predicting and managing degenerative lumbar spine disease.

Objectives:

The aim of this study was to evaluate the potential of ML models in predicting and determining the appropriate treatment for degenerative lumbar spine disease. Additionally, we sought to explore how these models could assist neurosurgery and spinal surgery residents in the decision-making process during outpatient evaluations.

Methods:

Using retrospective data from 92 patients with degenerative lumbar spine disease, we identified 83 clinical and radiological parameters critical for preoperative evaluation. The models were trained on this dataset and validated using 20 new prospective cases. Assessment of treatment modalities—conservative management, surgical decompression, or decompression with fusion—were conducted by nine medical professionals, including four neuro-spinal surgeons and five neurosurgery residents.

Results:

RF and SVM models demonstrated superior accuracy, achieving 87.7%, significantly outperforming the KNN model. These models accurately predicted treatment modalities and outcomes, showing promise as decision-support tools.

Conclusions:

The RF and SVM models can complement clinical expertise and support training for spinal surgery residents, enhancing their learning in decision-making for degenerative lumbar spine disease. Future work will focus on expanding datasets, refining models, and exploring advanced ML techniques to optimize their application in education and patient care.

Key Words:

Machine Learning, Lumbar degenerative disc disease, Artificial Intelligence, Random Forest, K-Nearest Neighbor, Support Vector Machine, Treatment Recommendation, Clinical Decision Support, Data-Driven Healthcare

Key Message:

This study presents a machine learning model that accurately recommends appropriate treatments for lumbar degenerative disc disease, serving as a decision-support tool for spine surgeons. It also provides a valuable resource for resident doctors, enhancing their training and decision-making capabilities.

Introduction:

Degenerative lumbar spine disease, a leading cause of disability, represents a substantial burden on healthcare systems^[4]. Clinical decision making about optimal management strategies for people with degenerative lumbar disc disease is complex. Multiple clinical and radiological factors are considered by the clinician to decide on surgery or conservative approach. Given a clinical and imaging scenario, a surgeon would intuitively decide on the management strategy- either surgical decompression with or without fusion or conservative management- based on his knowledge and experience. Most often it is subjective and not based on an algorithm or flow chart. Residents and surgeons in early phases of their learning curve traditionally hone the skill of decision making by observing senior surgeons making decisions during their residency and apprenticeship.

The process of decision making can be amplified by technology. Machine Learning (ML), a subset of Artificial Intelligence, offers promising solutions by leveraging data-driven approaches to analyse complex patient datasets. ML models can learn from experienced surgeons' decisions on multiple clinical scenarios and then suggest optimal treatment modalities when presented with new data. We conducted a study to develop and validate ML models to predict treatment strategies for degenerative lumbar spine disease, aiming to support spinal surgeons and train resident doctors in decision-making processes.

Study Design:

We conducted a retrospective study using data from 92 patients treated at the Narayana Institute of Neurosciences, Narayana Health, Bengaluru. Ethical approval was obtained from the local ethics committee. Patient consent was waived off due to the anonymized nature of the study. The cohort included 59 male and 33 female patients, all aged 18 years or older. Below Table 1 summarizes the distribution of patients across age groups.

Table 1: Count of Patients by age group considered for developing Machine

Age Group	Count
20-30	9
30-40	21
40-50	25
50-60	15
60-70	15
70-80	7

The study analysed patients treated conservatively (18 cases), with discectomy (45 cases), and fusion surgery (29 cases). Cases of spinal instability, scoliosis, or non-lumbar issues (e.g., cervical or thoracic conditions) were excluded. Clinical outcomes, including the

Oswestry Disability Index (ODI) at two weeks and two months post-treatment, were recorded.

Data Collection and Feature Selection:

Key factors influencing treatment decisions were identified in consultation with five neurosurgeons from Narayana Institute of Neurosciences. These factors, categorized into radiological, clinical, and post-operative outcome parameters, were classified as nominal, ordinal, discrete, or continuous variables. The target variable was the predicted treatment type: conservative management, surgery- discectomy, and surgery- discectomy with fusion. Below Table 2 Provides the factors considered

Table 2: Factors Considered for Developing the Machine Learning Model

Variable Category	Variable Sub Category	List of Variables
Qualitative	Nominal	Obese, Tone, Bulk, Tobacco Usage, Sedentary Job, Sensory, Disk Degeneration, Ligamentum Flavum Hypertrophy, Facetal Fluid Effusion, Disc Osteophytes Presence
	Ordinal	Pain VAS Score, Oswestry Disability Index, Power Hip, Power Knee, Power Ankle, Power EHL, DTR Knee Jerk, DTR Ankle Jerk, Plantar, SLRT, PIVD, Foramen Stenosis, Listhesis, Facetal Hypertrophy, ODI Index
Quantitative	Discrete	Age, Weight, Symptomatic Period, Aggravation of Symptoms Period
	Continuous	Canal Diameter

Learning Model:

The dataset underwent preprocessing to handle missing values, encode categorical variables, and standardize continuous data. This step ensured compatibility with three ML algorithms: Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbours (KNN). The robust dataset allowed for accurate modelling and validation, forming the foundation for developing decision-support tools tailored to the needs of spinal surgeons and residents.

Artificial Intelligence-Based Prediction Model:

The machine learning model was developed using Python within the Narayana Health Network domain. The model was trained using 92 retrospective data points. Three machine learning models—Random Forest (RF), Support Vector Machine (SVM), and K-Nearest Neighbor (KNN)—were implemented. These models were trained on 80% of the dataset, with the remaining 20% used for validation. The models were evaluated based on their ability to predict:

1. The need for surgery versus conservative treatment.
 2. The specific surgical intervention (e.g., discectomy or fusion).
 3. Post-operative recovery outcomes.
- ***Random Forest (RF):*** This ensemble learning method creates multiple decision trees and aggregates their predictions for more robust results^{[16][17]}.

- **Support Vector Machine (SVM):** This model is used for binary classification, separating patients based on their likelihood of requiring surgery versus conservative management^{[16][17]}.
- **K-Nearest Neighbor (KNN):** A simple algorithm that classifies cases based on the most similar past cases in the dataset.

Below Figure 1 shows the Python Program Output showing Accuracy obtained from Various Machine Learning Models viz. Random Forest, KNN and SVM

Figure 1: Python Program Output showing Accuracy obtained from Various Machine Learning Models viz. Random Forest, KNN and SVM

```
PS C:\Users\hsvsa\projnew\proj\model> python3.12 train.py
Saving model...
RF model saved.
Saving model...
KNN model saved.
Saving model...
SVM model saved.
PS C:\Users\hsvsa\projnew\proj\model> python3.12 app.py
RF Accuracy: 0.8775510204081632
KNN Accuracy: 0.5714285714285714
SVM Accuracy: 0.8775510204081632
```

The optimal machine learning model was incorporated into a natively developed interactive web application. This web application, developed within the Narayana Health Network domain, allows spine surgeons to input data based on the prospective new patient record. Below Figure 2 shows the Web Application Developed to Input Data of Prospective Patients for Backend Model to Predict Treatment Type

Figure 2 : Web Application Developed to Input data of Prospective Patient for Backend Model to Predict Treatment Type



Narayana Health

Data Intake Form

This Application is to Predict Spinal Treatment required for the Patient based on Artificial Intelligence Model. Fill the below Clinical and Radiological Evaluation of patient to Predict Treatment Required. Hover over labels for description.

HFO TOBACCO USAGE

☐ yes

☐ no

SEDENTARY/STRESSFUL JOB

☐ yes

☐ no

PAIN VAS SCORE

OSWESTRY DISABILITY INDEX

☐ Mild

☐ Moderate

☐ Severe

TONE

☐ yes

☐ no

PIVD_L1 L2

PIVD_L2 L3

PIVD_L3 L4

PIVD_L4 L5

PIVD_L5 S1

CANAL LENGTH_L1 L2

CANAL LENGTH_L2 L3

CANAL LENGTH_L3 L4

CANAL LENGTH_L4 L5

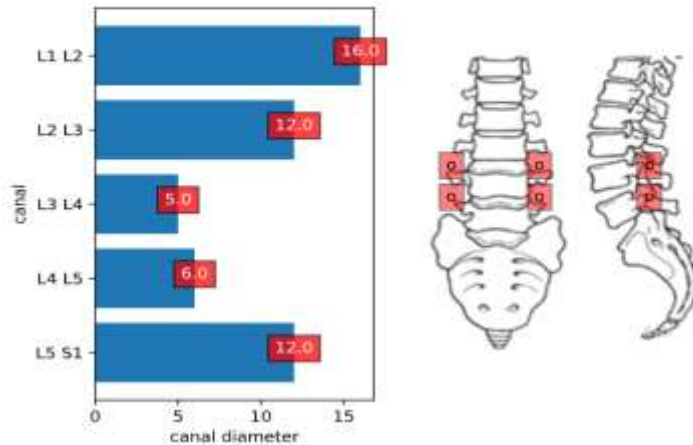
CANAL LENGTH_L5 S1

Results:

Based on the input of new patient prospective records, the machine learning model predicts the required treatment type. The output is presented in pictorial form, aiding spinal surgeons in making diagnostic decisions while providing a view of key factors. Below Figure 3 shows the Output from Machine Learning Model Predicting Treatment Type and Pictorially Representing Key Factors for Diagnostic Decisions

Figure 3 : Output from Machine Learning Model predicting Treatment type and Pictorially representing key factors for diagnosis decision

Predictions



knn

accuracy	0.5714285714285714
prediction	L4-L5 DISCECTOMY

rf

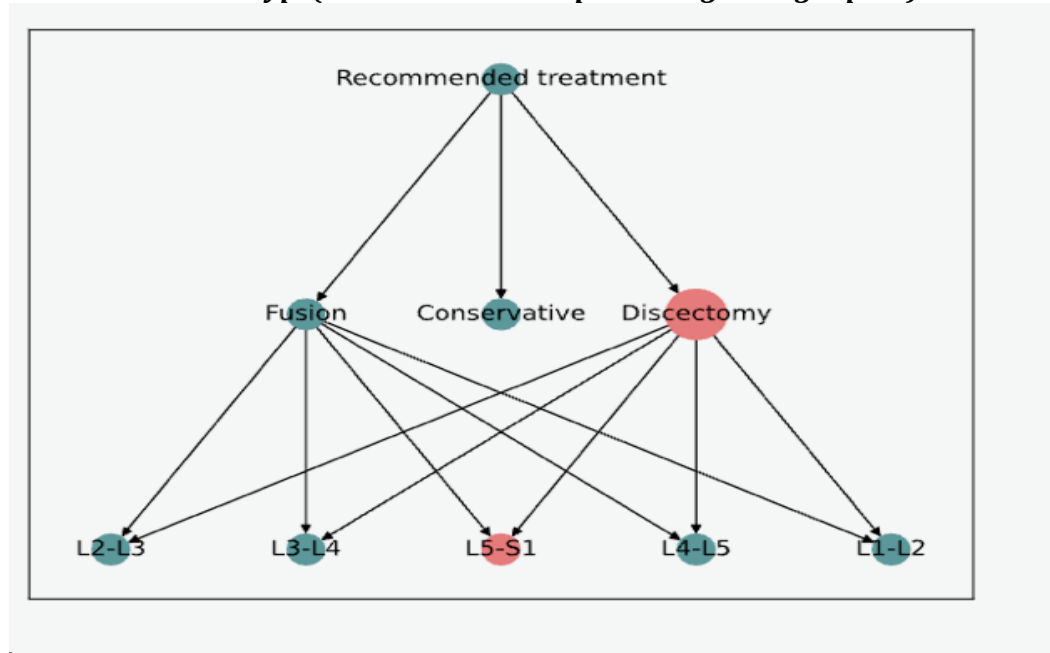
accuracy	0.8775510204081632
prediction	L4-L5 DISCECTOMY

svm

accuracy	0.8775510204081632
prediction	L3-L4 L4-L5 DISCECTOMY

Additionally, a network diagram was generated to represent the treatment type based on the model providing the highest accuracy. Below Figure 4 shows the Network Diagram from Machine Learning Model Representing Treatment Type (Red Color Nodes Representing the Right Path the Model predicted)

Figure 4: Network Diagram from Machine Learning Model representing the Treatment Type (Red Color Nodes Representing the right path)



To enhance the accuracy of the Machine Learning Model on continuous basis, an option was also provided in the portal for spinal surgeons to update the model by inputting new data records. This allows for continuous training of the machine learning model if the predicted treatment doesn't align with the actual diagnosis or if a new case is encountered. Below Figure 5 shows the Web Application to Input New Prospective Data for Machine Learning Model to Learn New Data Points

Figure 5: Web Application to Input new Prospective data for Machine Learning Model to learn new data points

Model Improvement

Please fill in details about your treatment.

What treatment was administered?

Form data is included.

Treatment data is included.

The machine learning model demonstrated high accuracy in predicting spinal disease types and recommending appropriate treatments. The evaluation metrics indicated that the model could reliably assist spinal surgeons in making informed decisions, thereby improving patient outcomes and providing a valuable resource for resident doctors, enhancing their training and decision-making capabilities.

Prediction Accuracy:

The machine learning models were evaluated using three different algorithms. The best performing model was then calibrated. Consistency between the model's predicted treatment type and the true observed treatment was noted. The accuracy observed for each model was as follows:

- Random Forest and SVM Model: 87.7% accuracy.
- KNN Model: 57.1% accuracy.

Based on these results, it was decided to proceed with the Random Forest and SVM models for prospective data prediction. Below Figure 6 shows the Accuracy of Various Machine Learning Models

Figure 6: Accuracy of various Machine Learning Model

RF Accuracy: 0.8775510204081632

KNN Accuracy: 0.5714285714285714

SVM Accuracy: 0.8775510204081632

Feature Importance:

Feature importance analysis in the RF model revealed that clinical and radiological factors, such as disc degeneration, foramen stenosis, and VAS pain scores, were the most predictive variables for determining surgical intervention and post-operative outcomes.

Methodology:

CRISP-ML Methodology Applied to Random Forest Classification

The CRISP-ML (Cross-Industry Standard Process for Machine Learning) methodology provides a structured approach to data science projects. The methodology's key steps aligned with the process followed in the Random Forest classification task:

1. **Process Understanding:**

- Objective: The primary objective was to classify medical conditions related to spinal treatments based on various patient features, enabling accurate prediction of the type of spinal treatment required.

2. **Data Understanding:**

- Data Exploration: The dataset contains features related to spinal conditions. Preliminary exploration, including understanding the feature types and their distributions, was conducted to prepare for the modeling phase.

3. **Data Preparation:**

- Feature and Label Selection: The dataset was split into features (X) and labels (Y). Features included various medical measurements, while the labels indicated the type of spinal treatment^[6-14].
- Data Splitting: The data was split into training and testing sets with a 70-30 ratio to ensure that the model had sufficient data to learn from while retaining a portion for evaluation.

4. **Modeling:**

- Model Selection: A RandomForestClassifier from the sklearn library was chosen due to its robustness in handling complex datasets with multiple features.
- Model Training: The Random Forest model was trained (fit) on the training data (X_train, Y_train), learning patterns that distinguish between different types of spinal treatments.

5. **Evaluation:**

Below Table 3 provides the Evaluation Parameters for Machine Learning Model Developed

Table 3: Evaluation Parameters for Machine Learning Model Developed

Evaluation Parameters	Algorithm	
	Random Forest Classification (RFC)	K- Nearest Neighbor (KNN)
Accuracy	87.7% indicates a high level of correct predictions from all test samples. The model is performing well.	57.1% indicates a significantly lower level of correct predictions. The model is underperforming compared to RFC.
Model Complexity	Complexity is high, due to the ensemble of decision trees. Advantages: Handles complex relationships and interactions in the data. Less prone to overfitting compared to single decision trees.	Complexity is low, as distance metrics are used to classify new instances. Advantages: Simple and intuitive, but performance heavily relies on the choice of k and can degrade with high-dimensional data.

Scalability	Scales are better with larger datasets due to the parallel processing of decision trees. Handles high-dimensional data effectively	Can be slow with large datasets, as it needs to compute distances to all training instances for every prediction. Performance can degrade with a large number of features.
Feature Importance	Provides feature importance scores, which can be useful for understanding which features are contributing the most to the predictions.	Does not provide feature importance directly. It treats all features equally without weighting them.

Evaluation Summary:

The Random Forest Classifier outperforms K-Nearest Neighbor by achieving a high accuracy of 87.7%. It is more robust for complex classification tasks and handles high-dimensional data better. Additionally, it provides feature importance, which can be insightful. On the other hand, KNN shows lower accuracy (57.1%) and may not be as effective for this dataset, especially with high-dimensional data.

6. Deployment:

The trained model is deployed to predict spinal treatment types for new patients. The model's high accuracy indicates it can be reliably used in a clinical setting to assist spinal surgeons in decision-making.

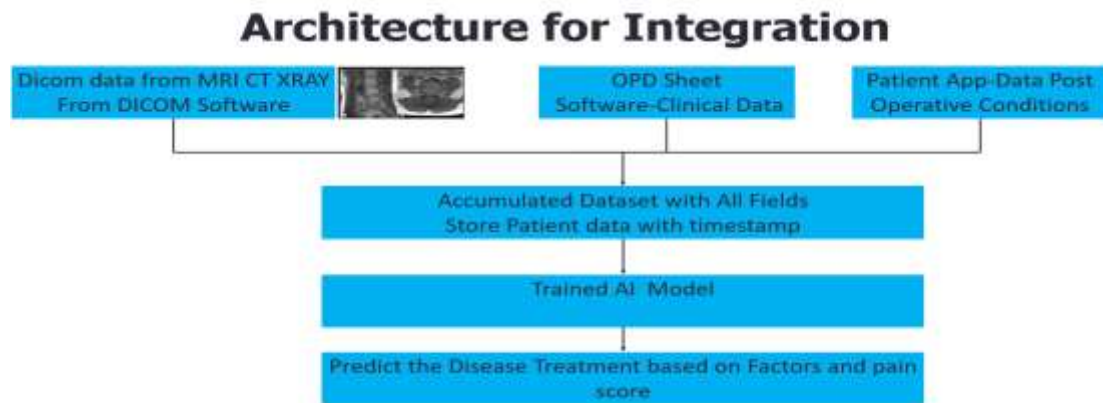
Limitations and Areas for Further Improvement in the Machine Learning Model:

The following areas of improvement were identified for future enhancements:

- The Oswestry Low Back Disability Index was not captured for retrospective data prior to the start of treatment.
- The prediction model does not consider instability and scoliosis cases.
- The machine learning model is based on only 92 retrospective data points. The accuracy of the model will improve as more data points are added and the model is further trained.
- Currently, only MRI data has been considered; inputs from X-ray and CT scan have not been included.
- Clinical data from nerve conductivity tests, such as EMG and ENMG, have not been considered.
- Further recommendations for conservative treatment (e.g., rest, nerve root block, or physiotherapy) and the type of physiotherapy are not captured.

To enhance the productivity of spinal surgeons, the ease of usage and integration of radiological and clinical data should be improved. This would eliminate the need for surgeons to manually input data and provide a comprehensive view to enable them to diagnose more preoperative cases. A schematic representation of such an architecture was prepared. Below Figure 7 shows an Architecture Model for Future Development to Integrate DICOM Software, OPD Sheet Clinical Data Software, and Patient App on Post-Operative Conditions

Figure 7: Architecture Model for future development to Integrate DICOM Software, OPD Sheet Clinical Data Software and Patient App on Post Operative Conditions



Compliance with Ethical Standards:

- Conflict of Interest: The authors declare that they have no conflict of interest.
- Ethics Committee Approval: Ethics approval was obtained from the Ethics Committee of Narayana Health, Bangalore, India. This study did not involve any new medicine, vaccines, devices, herbal remedies, or procedures.
- Patient Consent: Patient consent was not required as patient identified were not disclosed or compromised and was waived off by Ethics Committee due to anonymized nature of the study
- Financial Support and Sponsorship: None.

Discussion:

• **AI and Machine Learning in Spinal Surgery:**

The application of AI in spinal surgery, though relatively new, holds immense potential. The accuracy achieved by the Random Forest and SVM models demonstrates the ability of machine learning to support clinical decision-making. This study aligns with previous research, confirming that AI can assist in predicting spinal surgery outcomes with high accuracy.

• **Comparison of Machine Learning Algorithms^[13]:**

Random Forest, a robust ensemble method, outperformed KNN, an algorithm more suited for smaller datasets. SVM, effective for binary classification tasks, performed similarly to RF in this dataset, demonstrating its utility in clinical prediction tasks.

• **Limited Memory AI in Clinical Practice:**

The models used in this study represent Limited Memory AI, which utilizes historical data to make informed decisions. Feedback from real-world clinical applications allows these models to adapt to evolving clinical scenarios, a key feature of Limited Memory AI.

• **Challenges and Future Directions:**

While AI offers exciting prospects, challenges remain, including the need for large, diverse datasets, concerns about algorithmic bias, and integration into clinical workflows. Future studies should focus on refining these models and validating them in larger, multicentric datasets.

• **Learning from Other Healthcare Segments**

Some of AI Applications in Other Healthcare segments which have gained popularity are detection and diagnosis^[3], treatment, as well as outcome prediction^[4] and prognosis evaluation. Some of examples are AI Solutions for cervical cancer screening^[5], AI solution to identify eye disease signs^[7], AI Solutions for stomach cancer detection^[6] and Accurate diagnosis of pediatric diseases^[8]. Also Enthusiasm for applications of AI in healthcare has continued to grow, with early successes involving Chatbots, diagnostic tools, and radiological image analysis^[9]. Such studies can be of good Input for AI Application in spinal cases. AI can provide valuable support in clinical decision-making, surgical planning, and postoperative care, thus reducing variability^[10] in outcomes and enhancing surgical precision^{[11] [12]}.

Conclusion:

The CRISP-DM methodology applied to Random Forest classification led to the successful development of a predictive model with high accuracy. This model can enhance clinical decision-making, potentially improving patient outcomes in spinal treatments. AI, particularly ML models like Random Forest and SVM, holds significant promise in improving outcomes for spinal surgery patients. By offering personalized treatment recommendations based on clinical and radiological data, AI can support clinicians in making informed decisions and optimizing surgical outcomes.

Further research and validation are needed to fully integrate AI into routine spinal surgery practices. This model also represents a significant step towards addressing the shortage of spine surgeons in India. By providing data-driven insights, it can reduce spinal surgery failure rates and improve overall patient care. Future work will focus on expanding the dataset, enhancing usability, integrating data, improving model accuracy, and exploring additional machine learning techniques.

Appendix

Factors considered for Machine Learning Model with definitions are listed in Below Table 4

Table 4: Definition of Various Factors considered for developing Machine Learning Model

Field	Definition
PID	Patient ID
AGE	Patient Age
WEIGHT	Weight of Patient
OBESE	Obesity was considered as. a BMI greater than or equal to 30.
H/O TOBACCO USAGE	Patient Tobacco Usage
SEDENTARY/STRESSFUL JOB	Work that involves sitting down for most of the day. Any standing or walking is limited to two hours
PAIN VAS SCORE	Patients pain measured on a scale of 10cmno pain (0–4 mm), mild pain(5–44 mm), moderate pain (45–74 mm), and severe pain (75–100 mm)Worst Possible Pain-10

OSWESTRY DISABILITY INDEX	OSWESTRY DISABILITY INDEX ^[18] 1 No Disability or Mild disability- Patient can cope with most living or experiences pain with sitting, lifting and standing, 2-Moderate Disability-Pain remains main problem and daily living affected, 3-Severe disability-Pain Impinges on all aspects of patient life and patient is bed bound
TONE	Tone-resistance to passive stretch; The tension in the relaxed muscle” or “the resistance, observed
BULK	Bulk is the maximum thickness of muscle group measured at anatomically defined points.
SYMPTOMATIC PERIOD	Total duration of symptoms prior to intervention measured of Months
AGGREVATION OF SYMPTOMS	Number of Months back from when the Symptoms have aggravated. Treat 1 to 4 week as 1 Month
MUSCLE POWER	Assessment of Patients muscle strength according to MRC Grading ^[19] 0 - complete paralysis 1 - flicker of contraction possible 2 - movement is possible when gravity is excluded 3 - movement is possible against gravity 4 - movement is possible against gravity with moderate resistance 4+- movement is possible against gravity with moderate to strong resistance 5 - normal power-movement against gravity with strong resistance
DEEP TENDON REFLEX ^[20]	Deep Tendon reflex if totally absent is absent; If Reflex is small, less than normal, or a response brought out only with reinforcement is treated as Sluggish; Reflexes normal is treated as Normal
SENSORY	Hypoesthesia-any amount reduced perception of a particular deramatome to mechanical stimuli
STRATIGHT LEG RISING TEST ^[21]	If the patient experiences pain radiating down the leg (radiculopathy), when the straight leg is at an angle of between 30 and 70 degrees, then the test is positive and a herniated disk is a possible cause of the pain. Less than 30 and more than 70 degrees is treated as InSignificant
DEGENERATIVE DISC CHANGES ^[15]	According to Pfirrmann grading ^[22] system Grade 3 to Grade 5 was considered significant
POSTERIOR INTERVERTEBRAL DISC PROLAPSE ^[15]	Normal is N, Bulge- Diffuse outpouching of disc/annulus

	Whereas protrusion, extrusion and sequestration were all grouped together as significant changes.
CANAL LENGTH_	Diameter of Vertebral canal at particular disc level
NEURAL FORAMINAL STENOSIS	According to LEE system ^[23] of classification Grade 0 and Grade 1 was considered normal/acceptable and Grade 3 to Grade 5 were considered to be significantly affected
LISTHESIS GRADE	Ratio of the overhanging part of the superior vertebral body to the anteroposterior length of the adjacent inferior vertebral body. According to Mayerding classification system ^[24] Grade 0 and Grade 1 were considered normal/insignificant. Grade 3 to Grade 5 were considered as significant pathology.
LIGAMENTUM FLAVUM HYPERTROPHY	Enlargement or thickening of the Ligamentum flavum in the spine of size more than 4 mm was considered as significant issue.
FACETAL HYPERTROPHY	According to Weishaup system ^[25] of classification Grade 2 and Grade 3 were considered to be significant degeneration.
DISC OSTEOPHYTES	Series of bone spurs develop on multiple vertebrae, compressing the intervertebral discs and exerting pressure on the nerve roots or even the spinal cord. If found was considered significant.

References

1. Fei Jiang, Yong Jiang, Hui Zhi, Yi Dong, Hao Li, Sufeng Ma, Yilong Wang, Qiang Dong, Haipeng Shen, Yongjun Wang. Artificial intelligence in healthcare: past, present and future: Stroke and Vascular Neurology. BMJ Journals. 2017; 2(4), 230–243. DOI: <https://doi.org/10.1136/svn-2017-000101>
2. Rushikesh S Joshi, Darryl Lau, Christopher P Ames. Machine learning in spine surgery: Predictive analytics, imaging applications and next steps, Seminars in Spine Surgery. Elsevier. 2021; 33(2), 1–8. DOI: <https://doi.org/10.1016/j.semss.2021.100878>
3. Mohamed Khalifa, Mona Albadawy. AI in diagnostic imaging: Revolutionising accuracy and efficiency, Computer Methods and Programs in Biomedicine Update. Elsevier. 2024; 5, 1–12. DOI: <https://doi.org/10.1016/j.cmpbup.2024.100146>.
4. Monika Halicka, Martin Wilby, Rui Duarte, Christopher Brown. Predicting patient-reported outcomes following lumbar spine surgery: development and external validation of multivariable prediction models. BMC musculoskeletal disorders. 2023; 24(1), 1–16. DOI: <https://doi.org/10.1186/s12891-023-06446-2>
5. Hernán Darío Vargas-Cardona, Mérida Rodríguez-Lopez, Marcela Arrivillaga, Carlos Vergara-Sanchez, Juan P García-Cifuentes, Paula C Bermúdez, Andres Jaramillo-Botero. Artificial intelligence for cervical cancer screening: Scoping review. The International Journal of Gynecology & Obstetrics. 2023; 165(2), 566–578. DOI: 10.1002/ijgo.15179

6. Yiqian Zha, Cuili Xue, Yanlei Liu, Jian Ni, Jesus M De La Fuente, Daxiang Cui. Artificial intelligence in theranostics of gastric cancer, a review. *De Gruyter*. 2021; 3(3), 214–229. DOI: <https://doi.org/10.1515/mr-2022-0042>
7. Lorenz Adlung, Yotam Cohen, Uria Mor, Eran Elinav. Machine learning in clinical decision making. *Cell Press*. 2021; 2(6), 642–665. DOI: <https://doi.org/10.1016/j.medj.2021.04.006>
8. Li Ya-Wen, Liu Fang, Zhang Tian-Nan, Xu Fang, Gao Yu-Chen, Wu Tian. Artificial intelligence in pediatrics. *Chinese medical journal*. 2020; 133(3), 358–360. DOI: <https://doi.org/10.1097/CM9.0000000000000563>
9. Jordan P Richardson, Cambray Smith, Susan Curtis, Sara Watson, Xuan Zhu, Barbara Barry, Richard R Sharp. Patient apprehensions about the use of artificial intelligence in healthcare. *npj Digit. Med*. 2021; 4(1), 1–6. DOI: <https://doi.org/10.1038/s41746-021-00509-1>
10. Wireko Andrew Awuah, Favour Tope Adebusoye, Jack Wellington, Lian David, Abdus Salam, Amanda Leong Weng Yee, Edouard Lansiaux, Rohan Yarlagaadda, Tulika Garg, Toufik Abdul-Rahman, Jacob Kalmanovich, Goshen David Miteu, Mrinmoy Kundu, Nikitina Iryna Mykolaivna. Recent Outcomes and Challenges of Artificial Intelligence, Machine Learning, and Deep Learning in Neurosurgery. *World Neurosurgery*. 2024; 23, 1–9. DOI: <https://doi.org/10.1016/j.wnsx.2024.100301>
11. Alowais S A, Alghamdi S S, Alsuhebany N, Alqahtani T, Alshaya A I, Almohareb S N, Aldairem A, Alrashed M, Bin Saleh K, Badreldin H A, Al Yami M S, Al Harbi S, Albekairy A M. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC medical education*. 2023; 23(1), 689. DOI: <https://doi.org/10.1186/s12909-023-04698-z>
12. Sushanta Kumar Das, Ramesh Kumari Dasgupta, Saumendu Deb Roy, Dibyendu Shil. AI in Indian healthcare: From roadmap to reality. *Intelligent Pharmacy*. 2024; 2(3), 329–334. DOI: <https://doi.org/10.1016/j.ipha.2024.02.005>
13. Uddin S, Khan A, Hossain M E, Moni M A. Comparing different supervised machine learning algorithms for disease prediction. *BMC medical informatics and decision making*. 2019; 19(1), 281. DOI: <https://doi.org/10.1186/s12911-019-1004-8>
14. Beatriz Remeseiro, Veronica Bolon-Canedo. A review of feature selection methods in medical applications. *Computers in Biology and Medicine*. 2019; 112, 1–9. DOI: <https://doi.org/10.1016/j.combiomed.2019.103375>
15. Kushchayev S V, Glushko T, Jarraya M et al. ABCs of the degenerative spine. *Insights Imaging*. 2018; 9, 253–274. DOI: <https://doi.org/10.1007/s13244-017-0584-z>
16. Jackins V, Vimal S, Kaliappan M et al. AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes. *J Supercomput*. 2021; 77, 5198–5219. DOI: <https://doi.org/10.1007/s11227-020-03481-x>
17. Mahajan P, Uddin S, Hajati F et al. A comparative evaluation of machine learning ensemble approaches for disease prediction using multiple datasets. *Health Technol*. 2024; 14, 597–613. DOI: <https://doi.org/10.1007/s12553-024-00835-w>
18. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)* PubMed. 2000; 25(22):2940–52. DOI: 10.1097/00007632-200011150-00017
19. Naqvi U, Sherman AL. Muscle Strength Grading. *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2025. <https://www.ncbi.nlm.nih.gov/books/NBK436008/>

20. H Kenneth Walker, W Dallas Hall, J Willis Hurst. Deep Tendon Reflexes. Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition. Boston: Butterworths.1990; Chapter 72. <https://www.ncbi.nlm.nih.gov/books/NBK396/>
21. Camino Willhuber GO, Piuizzi NS. Straight Leg Raise Test. StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025. <https://pubmed.ncbi.nlm.nih.gov/30969539/>
22. Urrutia J, Besa P, Campos M, Cikutovic P, Cabezon M, Molina M, Cruz JP. The Pfirrmann classification of lumbar intervertebral disc degeneration: an independent inter- and intra-observer agreement assessment. Eur Spine J. 2016; 25(9):2728-33. DOI: 10.1007/s00586-016-4438-z
23. Lee S, Lee JW, Yeom JS, Kim KJ, Kim HJ, Chung SK, Kang HS. A practical MRI grading system for lumbar foraminal stenosis. AJR American Journal for Roentgenology. 2010;194(4):1095-8. DOI: 10.2214/AJR.09.2772
24. Koslosky E, Gendelberg D. Classification in Brief: The Meyerding Classification System of Spondylolisthesis. Clin Orthop Relat Res. 2020; 478(5), 1125-1130. DOI: 10.1097/CORR.0000000000001153
25. Sriselvakumar S, Worsley C, Feger J, et al. Weishaupt grading system of lumbar facet joint degeneration. Radiopaedia.org.2025. DOI: <https://doi.org/10.53347/rID-158404>