Frontiers in Health Informatics
ISSN-Online: 2676-7104
www.healthinformaticsjournal.com
2024; Vol 13: Issue 3
Open Access

Ensemble-Based Orthopedic Biomechanical Characteristics Prediction Using Probabilistic Reasoning And Machine Learning Approach

Dr. G. Sasipriya[1] and Dr.B.Lavanya[2], *
[2]Research Scholar, Department of Computer Science, University of Madras, Chennai, India. E-Mail: lavanmu@gmail.com
[1]Associate Professor, Department of Computer Science, University of Madras, Chennai, India. E-Mail: msasipriyam@gmail.com

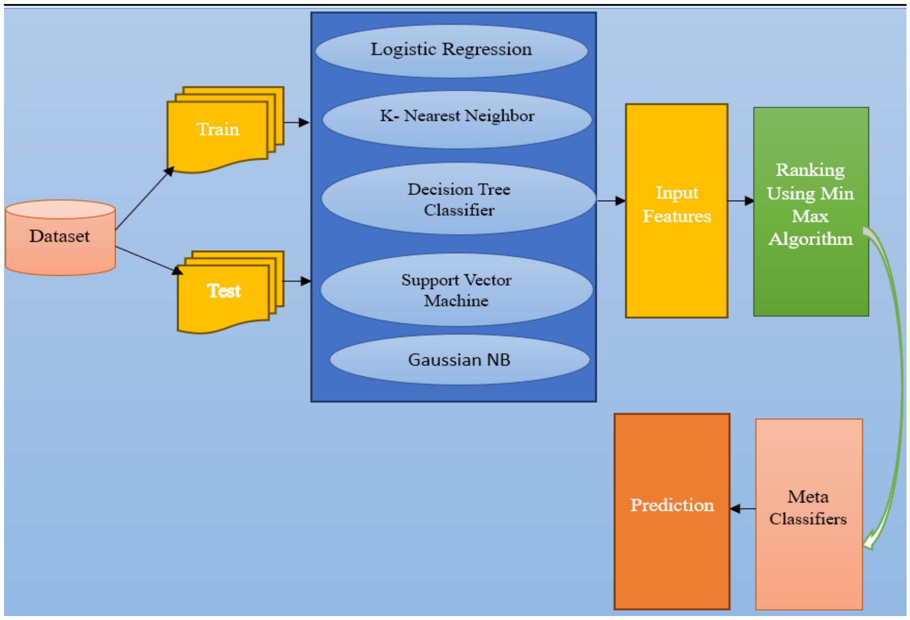
Cite this paper as: G. Sasipriya, B.Lavanya (2024) Ensemble-Based Orthopedic Biomechanical Characteristics Prediction Using Probabilistic Reasoning And Machine Learning Approach. *Frontiers in Health Informatics, 13* (3), 10112-10124

Abstract
In hospitals, medical data are growing increasingly complicated and heterogeneous in significant data circumstances. Massive medical data management requirements have yet to be satisfactorily addressed by the conventional way of manual computation. With the advancement of artificial intelligence and machine learning, this medical diagnosis model has been created. Now withstanding to enhance the perforation in the additional determination of orthopedic disease data. The purpose of this paper is to provide a machine-learning technique for auxiliary categorization prediction that may be used to aid in the diagnosis of orthopedic diseases. This study shows how to build ensembles of diverse classifiers by stacking many types of classifiers, including the Gaussian NB, logistic regression, decision tree regressor, K-Nearest Neighbor algorithm (KNN), and Support Vector Machines (SVM). Five different base classifiers are used, and min-max ranking is used to give weightage to various base classifiers. This work is crucial because the proposed algorithms can make quick judgments on diagnoses of orthopedics with a good level of accuracy. The ensemble stacking, the ensemble stacking with equal weightage, and the ensemble stacking giving rank by utilizing the min-max method are all proposed in this study.

- *The findings of this research indicate that stacking, followed by ranking, using the min-max method can be used in developing expert systems that are both efficient and effective in diagnosing disc hernia and spondylolisthesis.*
- *This study suggests an orthopedic diagnosis classification and prediction model with an accuracy of 97.80% and 97.85%.*
- *The proposed system aims to reduce medical staff labor, helps patients prevent and recover early, and provides real supplemental clinical attention.*

Keywords
Medical data, Machine Learning, Ensemble Stacking, KNN, SVM

Graphical abstract



1. Introduction

The term "Orthopedics" was first used by French scholar Nicholas Andry in 1741 when he published Orthopedic. Orthopedic, sometimes known as orthopedic, is a branch of surgery that deals with musculoskeletal disorders and injuries. The musculoskeletal system is made up of nerves, ligaments, tendons, muscles, joints, and bones. Orthopedists in India have a history that goes back more than 4,000 years. Orthopedic doctors reduced fractures and performed therapeutic trepanations throughout antiquity, beginning with the Harappan culture [1]. In India, 30% of the population lives in poverty [2] [3]. When people have similar conditions, doctors will give them the same diagnosis and treatment plan, which will make the diagnosis less specific. Sometimes, because people are so uniquely distinct from one another, patients who have been diagnosed with the same ailment and have been given the same treatment will have very varying levels of improvement. There is no question that this will result in a delay in providing patients with the most effective therapy possible, and it may even result in major medical incidents. Under these conditions, the medical establishment, which has access to vast amounts of health records, started investigating new ways of conducting business to improve day-to-day operations by gathering and analyzing data [4] [5].

At this time, electron microscopy, pharmaceutical research, and supplementary diagnostics are some of the areas in which the pharmaceutical and healthcare industries have benefited from the application of machine learning and artificial intelligence techniques. Despite the demand, the implementation of AI into clinical practice is still in its preliminary stages. Patients can receive medical care that is both more accurate and more intelligent when medical professionals and computing systems work together [6]. As a result of this, one of the most exciting areas of research within the field of intelligent medicine right now is focused on developing a classification and prediction model of auxiliary diagnosis that is predicated on healthcare trials. AI chooses the most relevant patient-specific imaging examination and generates the most appropriate protocol by incorporating information from the patient's medical records, which may include symptoms, test results, and findings from a physical examination [7][8].

This paper explains how to distinguish between a disc hernia and spondylolisthesis by analyzing biomechanical parameters derived from the pelvic and lumbar spine structure and orientation. Pelvic incidence, pelvic tilt, lumbar lordosis angle, sacral slope, pelvic radius, and spondylolisthesis grade are all taken into account as biomechanical factors. Biomechanical features are mapped to clinical diagnoses such as disc herniation, spondylolisthesis, or normality, and artificial neural networks are utilized to learn this mapping [9]. On the other hand, spondylolisthesis is a disorder in which one of the vertebrae slides forward out of position about the vertebrae below it. This can cause back pain and other symptoms. There are two different kinds of spondylolisthesis, namely spondylolysis and degenerative. In most cases, degenerative spondylolisthesis is brought on by age. Normal wear and tear on the intervertebral discs can cause their physical properties to shift with time. In addition, discs may get shorter, less flexible, or even larger as time passes. Because the vertebra is out of place, there is a possibility of a constriction along the spinal channel, which would put pressure on the nerves. Patients who are over the age of 50 are typically the ones who are diagnosed with degenerative spondylolisthesis.

In this study [10], machine learning models are developed to predict recurrent lumbar disc herniation (rLDH) following percutaneous endoscopic lumbar discectomy. rLDH stands for recurrent lumbar disc herniation (PELD). They were able to accomplish this by identifying critical variables for predicting rLDH such as higher BMI, lower FO, Modic alterations, disc calcification in a non-protrusive location, and herniation type (non-contained herniation). This research presents a quick biomechanical and neural network-based disc hernia and spondylolisthesis detection technique. In this work, trained neural networks classify new cases as disc hernia, spondylolisthesis, or normal. Disk hernia and spondylolisthesis have similar symptoms, therefore misdiagnosis is common. The proposed technologies can quickly and accurately make such complex diagnoses. Public database data trains feedforward and radial basis function networks. This study suggests that neural networks can be used as expert systems to diagnose disc hernia and spondylolisthesis [11].

The technique makes use of biomechanical characteristics in addition to ensemble stacking with a weighted ensemble minimax ranking model. It is now possible to diagnose fresh scenarios by utilizing a trained stacking model in conjunction with the min-max ranking approach. This is made possible as a consequence of the many different kinds of procedures that are discussed in this research study. Patients are classified into one of three groups according to whether or not they have a herniated disc, spondylolisthesis, or normal. This classification is based on the severity of the patient’s condition. The associated works of orthopedics, the methodology, the proposed model, the findings, and the dispute are all going to be described in the next section of this article.

2. Machine Learning Methods

2.1. Decision Tree Classifier

It is an algorithm for learning through supervision. In this particular technique, the data are continuously broken down into more manageable portions until they reach their class. It makes use of concepts like nodes, edges, and leaf nodes, among others. The entropy of our resource is the primary thing we calculate when using the Decision Tree classifier. It gives us an indication of the degree of uncertainty associated with our database. The uncertainty value should be as low as possible for the best possible categorization results. The information gained associated with each feature is computed. After spitting the database, this then shows us the degree to which the amount of uncertainty has decreased. After finally computing the total information gain for every feature, the next step is to determine which parts of the database have the highest information gains. The procedure is repeated numerous times until all of the nodes have been eliminated [12].

2.2. SVM

The Support Vector Machine (SVM) is an example of a supervised learning algorithm that determines which hyperplane will split the dataset most effectively. The following are definitions of the two most important terms that will appear frequently in this text. The locations that are located in the closest proximity to the hyperplane are referred to as support vectors. A hyperplane is a subspace that has a dimension that is one less than the dimension of its surrounding space. It provides the purpose of separating the room into several distinct portions [13]. If we start with a space that has three dimensions, the succeeding hyperplane will be a plane that only has two dimensions. Similarly, the hyperplane of a plane with two dimensions would be a line with just one dimension. Margin refers to the distance that exists between the hyperplane and the data point that is located closest to it on each side. The term ”kernel” refers to a mathematical function that is used to convert one set of data into another format. Linear, nonlinear, polynomial, and other types of kernel functions are frequently used. You can see the hyperplane as a linear, one-dimensional line that separates the data if you are performing a straightforward classification problem using just two features [14].

2.3. Logistic regression

Logistic regression is a commonly utilized technique in the realm of machine learning. It is commonly used for tasks involving binary classification. This statement aims to predict the probability of an observation being categorized into a given group. This statement usually results in binary outcomes, like true or false, or yes or no. Logistic regression outcomes consist of probabilities ranging from 0 to 1, unlike the continuous measures provided by linear regression forecasts. For the algorithm to function properly, it must precisely represent the relationship between the target variable's probability and the input features. The logistic function, often known as the sigmoid function, is used to constrain projected probabilities between zero and one. Logistic regression is a statistical method used to predict the chance of an observation being categorized into a specific category. This process involves initially determining the coefficients associated with each input feature, which are then utilized to compute the probability. During the training phase, the model employs a range of optimization techniques, such as gradient descent and maximum likelihood estimation, to derive these coefficients from the training data. The goal is to minimize the difference between the predicted probabilities and the actual class labels in the training data. The logistic regression model can be trained to make predictions about the probability that new data will be classified into a given set.

The probabilities are then compared to a set threshold, which determines the creation of binary classifications. If the expected likelihood is above 0.5, the observation is categorized into one class; otherwise, it is categorized into the other class. Logistic regression offers several advantages, such as the ability to deal with both numerical and categorical input features, as well as the straightforwardness and comprehensibility of the approach. However, it assumes that there is a direct relationship between the characteristics of the input and the logarithmic probabilities of the outcome, which may not always hold true in practical situations. However, despite this fact, logistic regression remains a widely used and successful method for binary classification tasks in multiple disciplines [15].

2.4. K-nearest neighbor

According to the KNN algorithm, related objects tend to cluster together, complementary elements tend to cluster together. 1. Data loading 2. Choose a starting value for K that is equal to the number of close neighbors you want to start 3. Every instance in the data 3.1 Determine the data-driven separation between the query example and the current instance. 3.2 Complement an ordered set with the length and the number of the example. 4. using the distances, arrange the sorted list of distances and indices from smallest to largest. 5. In this step, select the first K items from the sorted set. 6. Obtain the names of the K-selected entries in step 6. 7. If a regression analysis is performed, the average of the K labels should be returned 8. Return the middle K labels if classifying 8.

2.5. Na¨ıve Bayes

One example of a probabilistic classifier is the Naive Bayes algorithm for classification. It is founded on probability models that involve stringent assumptions of independence. Quite frequently, the independence assumptions do not have any effect on the real world. Because of this, others view them as being naive. The naive Bayes classifier is one of the simplest techniques for the classification task that is nonetheless capable of achieving decent accuracy. It was developed by the mathematician and statistician George Bayes. Although it is not always able to compete with algorithms that are much more refined, such as decision trees, there are some application domains in which it does not lag too far behind and may even be superior. Text classification is the most prominent example of this. While it is true that in many situations it can't compete with such algorithms, there are times when it does not lag very far behind. Because of its conceptual, implementational, and computational simplicity, testing it alongside or before classifiers with a higher level of sophistication is both simple and economical.

3. Method Details

The proposed method of this paper is represented in Figure 1, which tells about the dataset, machine learning algorithms, ranking using min-max, meta classifier, and prediction of the ortho diagnosis data.

3.1. Data Analysis

The dataset we obtained from the public database, is about the biomechanical feature of orthopedic patients. The pelvic incidence, pelvic tilt, lumbar lordosis angle, sacral slope, pelvic radius, and stage of spondylolisthesis are the ligamentous parameters that are taken into consideration. In addition, this paper discusses two distinct diagnostic approaches, which are outlined in the following paragraphs. The ortho diagnosis data-2 this system is provided with the six biomechanical features attributes that were mentioned previously, and as a result, it concludes as to whether the patient has spondylolisthesis, disc hernia, or whether the patient is normal. The ortho diagnosis data-1 with two classes, the system helps diagnose spondylolisthesis and disc hernia in patients who have already been determined to be suffering from either of these conditions. The decision-making process likewise uses the same six biomechanical characteristics as the evaluation process [11].

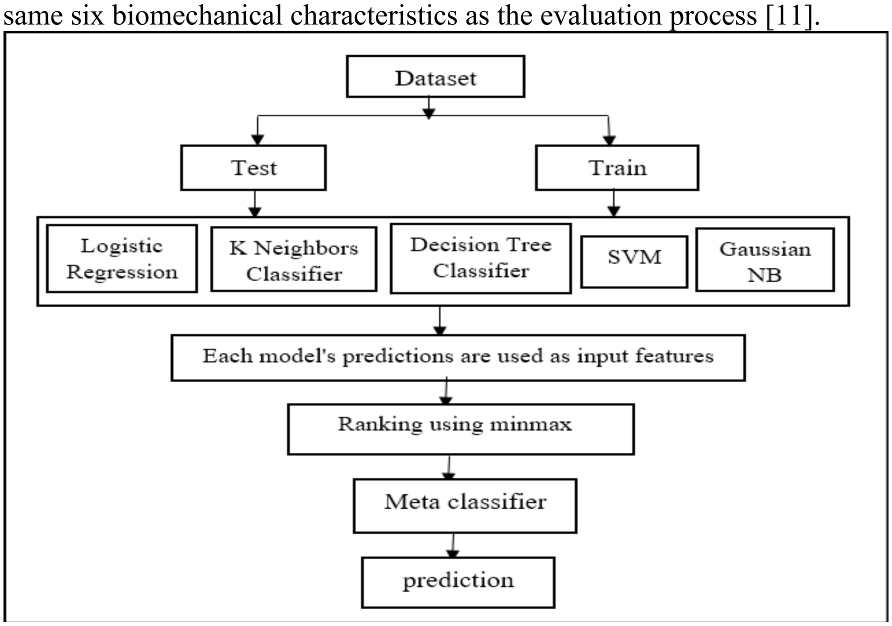


Figure 1: Proposed ensemble stacking using min-max ranking

3.2. Stacking

Machine learning ensemble approaches are the cutting-edge technology that arose to address these issues. Generally speaking, there are three types of ensemble methods: bagging, boosting, and stacking. Bagging is a technique for permitting redundancy in the learning process by randomly classifying training data to a prediction model. Similarly boosting works to bagging, except it gives more weight to the data that the prior model failed to categorize correctly. Stacking is a super-learning ensemble method. To produce fresh data from expected results, it employs several different models as base learners, including random forest, k-nearest neighbors, and support vector machine [16]. The meta-learner, an additional predictive model, is then employed to get the predicted value from the updated dataset. The field of ensemble approaches has been expanding at a steady rate, with many studies focusing on bagging and boosting ensembles. If bagging and boosting ensemble approaches are used to build strong predictive models utilizing the same incomplete and weak prediction models, then staging ensemble techniques are available to integrate different predictive models [17].

A model that makes use of the predictions generated by numerous machine learning models to generate a predictive model that is even more accurate is known as an ensemble of models. Through the use of numerous machine learning models and an ensemble of models, we attempted to maximize the predictions made by our models by giving weights to each model based on how well it performed. An algorithm for ensemble machine learning that is capable of discovering how to optimally mix each of the models that comprise an ensemble to achieve optimal results. The primary goal of a typical machine learning model is to generate a relationship function, which acts as a mapping function between input and output. Stacking performs an action that is one step farther than the ordinary by learning the relationship between the estimation result from each of the ensembled models including our prediction models and the actual outcome. This allows it to function on a level that is one step above the ordinary. A stacked ensemble's fundamental structure is made up of two or more level-0 base models and a higher-level meta-model (level-1 model), each of which fulfills one of the following roles in the ensemble's overall operation:

- Base-Models is also known as Level-0 Models, are models that can predict out-of-sample data while still fitting the training data.
- Meta-model is also known as Level-1 Model, A model that fits on the prediction from lower-level base models and learns to combine the predictions in the most effective way possible.

3.2.1. Creating a stacking model

The first part of this article demonstrates that stacking is a two-level model, with a ”Level 0” model used to generate classification labels that become new features in the data and a ”Level 1” model used to generate the final label predictions. Both of these models are described in more detail in the following paragraphs. It has been determined that Level 0 should function as a dictionary for the categorization of machine learning algorithms. In this study, we decided to use five distinct kinds of models: logistic regression, decision tree classifier, support vector machine, Gaussian NB, and K-nearest neighbors.

The latter five are all high-performing classifiers, and the logistic regression adds some variance to the results. The selection of algorithms with high performance was done on purpose to test whether or not the stacking model can successfully outperform them. It has been decided to use a logistic regression for the ”Level 1” procedure also known as the ”final estimator”). The results of the experiments indicated that it was the individual method that performed the best on this dataset; hence, it was chosen to serve as the level 1 model to push higher performance from the stacking model. The fundamental procedures of any stacking method can be summed up as follows:

- K-Fold Cross-Validation is to split the dataset into K roughly equal parts. The dataset will be represented as D and the amount of folds as K. For each i ranging from 1 to K, the fold is represented as Di.
- The Meta-Model Training is performed From 1 to K, for every iteration i: The validation set will be fold Di. Use the remaining folds D\Di to train N distinct basis models, such as decision trees or support vector machines. Here, j ranges from 1 to N, and these models are represented as Mi,j.
- The Predictions Made Outside of the Sample, All base models Mi,j must be met. Find the results of folding Di using M i,j. These predictions are represented as ˆYi,j.
- Feature Construction - Merge all out-of-sample predictions ˆY i, j into a fresh dataset X′, where a data point is represented by each row and a prediction from a base model is represented by each column.
- Meta-Model Training, (such as logistic regression or random forest) using X′ and the associated true outcomes Y. The meta-model is represented as M ′.1

3.2.2. Creating a weightage using the Minmax algorithm

These algorithmic steps explain the weights given to the machine learning models by using the min-max algorithm.

Input: Dataset D= Dataset1 (D1), Dataset2 (D2)
$m_1, m_2, m_3, \ldots m_n$ - Base learning algorithm.
$p_1, p_2, p_3, \ldots p_n$ - Predictions.

- The initial datasets will be indicated as D_1 and D_2. As an example, we divide each dataset into two halves, D_1 and D_2, with D_1 representing the training set and D_2 representing the testing set.
- The $m_1, m_2, \ldots$, and m_n are the base learning algorithms. For every algorithm m_i, we acquire Predictions by training a model on D_{1train} and then evaluating its performance on D_{1test}.

In the same way, we achieve predictions $p_{2\text{test}}^{(i)}$ by training on D_{2train} and evaluating on D_{2test}.

- To make sure that the values of each model's predictions $p_{\text{test}}^{(i)}$ are between 0 and 1, the Min-Max scaling transformation is used. Here is one way to depict this transformation:

$$p_{\text{scaled}}^{(i)} = \frac{p_{\text{test}}^{(i)} - \min(p_{\text{test}}^{(i)})}{\max(p_{\text{test}}^{(i)}) - \min(p_{\text{test}}^{(i)})}$$

The text is being resized. The square root of P_{test} raised to the power of i minus the minimum of P_{test} raised to the power of i is equal to (i). the maximum value of the function $p_{\text{test}}^{(i)}$ – the minimum value of the function $p_{\text{test}}^{(i)})$.

- By utilizing the Min-Max ranking method, the scaled predictions of each model are rated. This uses the scaled value to rank all of the predictions.

- Lastly, a meta-classifier merges the ranked predictions from all models to produce the final prediction. Combining the ranking predictions using a weighted or unweighted method is the mathematical aspect of this process that depends on the meta-classifier utilized. As an illustration, the metaclassifier may merge the predictions of all n basic models in the following way:

$$\text{Final prediction} = \frac{1}{n}\sum_{i=1}^{n} \text{Rank}(p^{(i)}_{\text{scaled}})$$

- The formula for the final prediction is the sum of the rank of the function $(p^{(i)}_{\text{scaled}})$ raised to the power of n, where n is the number of elements in the input matrix i=1. the rank given to the scaled prediction of model i is represented by Rank $(p^{(i)}_{\text{scaled}})$. The average rank of all models would be used to make the final prediction.

4. Method Validation

This study describes an automated method for diagnosing spondylolisthesis and disc herniation based on biomechanical characteristics of the patient's spine. The ensemble stacking approach and stacking based on min-max ranking are two of the methods that are used to construct quick diagnostic systems that can diagnose new cases with a level of accuracy that is considered to be acceptable as shown in Tables 1 and 2 for both datasets. Patients can be classified as having disc hernia, spondylolisthesis, or normal when using the established ortho diagnostic 2 Ensemble stacking approach and stacking using min-max ranking-based expert systems. Tables 1,2 show the accuracy values of standalone and stacked ensemble classification models for both ortho diagnosis datasets 1 and 2. In the case of standalone models for dataset 1, the Decision tree classifier has an accuracy of 84.44 followed by the Logistic regression classifier with 83.65, the values for dataset 2 are 81.43 and 78.57 respectively. In the case of the stacked ensemble, the model with Logistic regression as level 1 (Meta learner) has the highest accuracy of 87.1 for dataset 1 and 84.26 for dataset 2.

Table 1: Ortho Diagnosis Data1 Accuracy

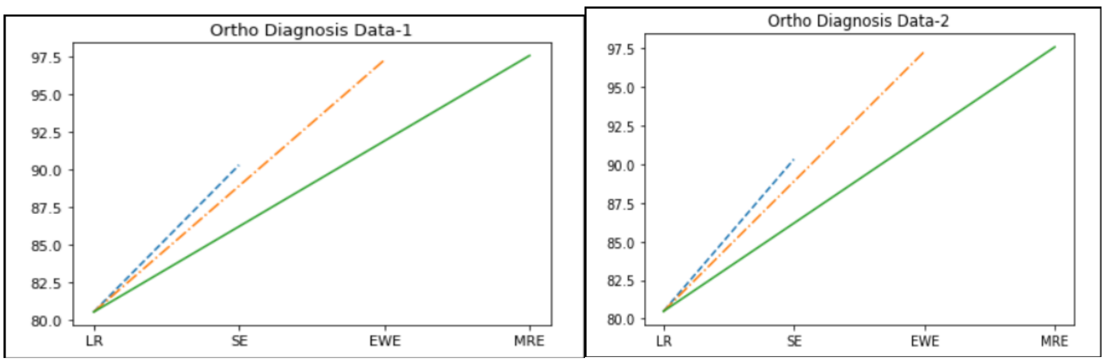
Models	Classifier (Single)	Stacked Ensemble
LR	83.65	87.1
KNN	71.98	83.87
DTC	83.89	75.81
SVM	79.92	83.87
BAYES	78.97	83.87

Table 2: Ortho Diagnosis Data1 Accuracy

Models	Classifier (Single)	Stacked Ensemble
LR	78.57	82.26
KNN	69.84	83.87
DTC	81.43	75.81
SVM	79.13	83.87

BAYES	79.76	83.62

When applied to new situations, it has been discovered that the Ensemble stacking approach as well as the suggested stacking utilizing min-max ranking both motivational yield performances. The best overall classification accuracies were obtained from the ensemble stacking method, which was 91.935%, and the stacking method that used min-max ranking, which was 97.850% has been visualized in Figures 3a and 3b. On the other hand, it should be pointed out that the stacking method that used min-max ranking was able to attain greater recognition rates on the training data. i.e., a higher capacity for generalization in comparison to the stacking of the ensemble. The equal weightage ensemble method for both datasets gives good results as shown in Figures 2a and 2b. In Figures, 2a and 2b EWE denotes the equal weightage ensemble method, and MRE represents the min-max ranking ensemble method.



(a) Figure 2a (b) Figure 2b

Figure 2: Stacked Ensemble, Equal weightage ensemble & Minmax ranking ensemble for diagnosis data-1&2

In addition, utilizing Ortho diagnostic 1 derived from the same public database, these technologies make it possible to categorize individuals as having either spondylolisthesis or disc hernia, depending on which condition they have. These systems are only applicable in circumstances in which a medical professional has first established the presence of spondylolisthesis or disc herniation as the underlying condition. The significance of the ortho diagnostic 1 system resides in the fact that it might be difficult to determine if a patient has a disc hernia or spondylolisthesis because both conditions present similar symptoms. This makes it difficult to decide which ailment the patient has. That is to say, it is somewhat easier to ascertain that a patient has either of the disorders (disc hernia or spondylolisthesis), as opposed to precisely determining which one of the two conditions a patient suffers from in their case. The best overall classification accuracies achieved for both datasets from ensemble stacking are 85.484% and 91.935, while the best overall classification accuracies obtained from proposed ensemble stacking using min-max ranking are 97.800% and 97.850%, respectively from Table 3.

Table 3: prediction of stacking and proposed stacking using min-max

S. No	Models	Ortho-diagnosis1	Ortho-diagnosis2
1.	Ensemble stacking model	85.484	91.935
2.	Proposed Ensemble stacking min-max ranking model	**97.800**	**97.850**
3.	XGBoost [4]	-	0.951%
4.	Radial basis function networks[11]	96.67%	88.39%,

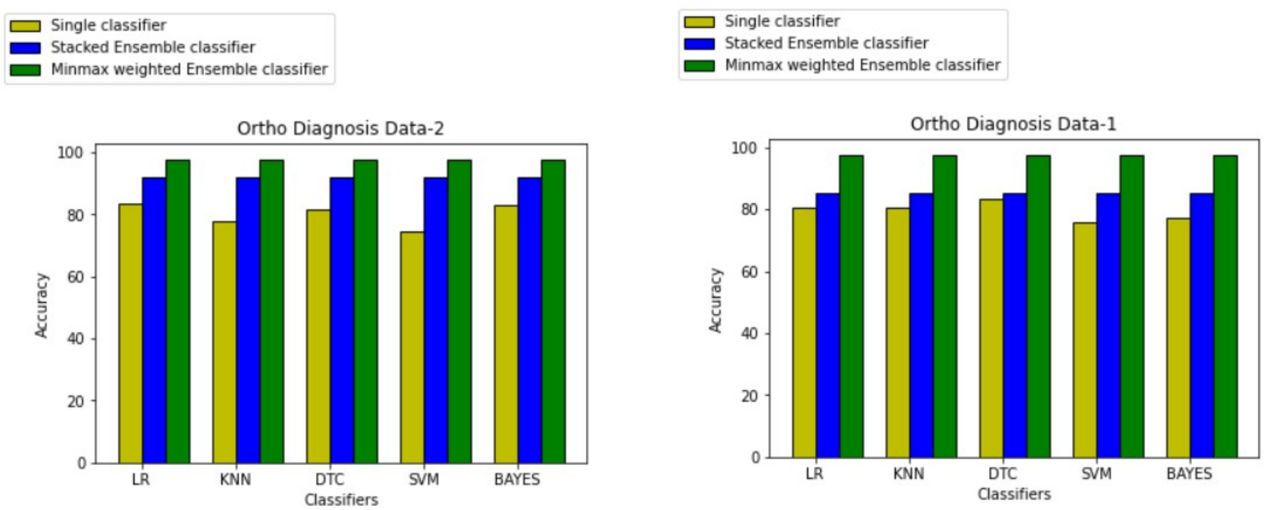


Figure 3: Performance comparison of dataset-1 & 2

4.1. Performance Validation Measures:

The RMSE values for the datasets are displayed in Tables 4 and 5. The root-mean-squared error (RMSE) quantifies how much a model's predictions deviate from the observed data. When comparing the model's performance to the actual data, a reduced root-mean-squared error (RMSE) indicates a more accurate prediction. When comparing the two datasets, it is clear that stacked ensemble models outperform standalone models. This provides more evidence that stacking, which combines the predictions of different models, improves performance concerning RMSE.

The stacked ensemble using a Support Vector Machine (SVM) as its base learner routinely outperforms all other models on both datasets. This shows that SVM greatly helps in lowering prediction errors when it's part of the stacked ensemble. On closer inspection, however, it becomes clear that, on these particular datasets, the stacked ensemble model utilizing Logistic Regression as the meta-learning model achieves superior performance. The logistic regression-based stacked ensemble outperforms SVM as a base learner, with even lower RMSE values, demonstrating better prediction accuracy.

The results show that the logistic regression-based meta-learning model outperforms SVM when it comes to merging the predictions of base learners to produce more accurate ensemble predictions for the provided datasets, even if SVM is great at capturing complicated relationships in the data. To conclude, the stacked ensemble model with Logistic Regression as the meta-learning model outperforms standalone models and other ensemble configurations for both datasets, as shown by the RMSE values in Tables 4 and 5. This approach effectively improves predictive performance.

Table 4: Root Mean Square Error Validation for Ortho Diagnosis Data1

Models	Classifier (Single)	Stacked Ensemble
LR	0.3479	0.1162
KNN	0.1466	0.1140
DTC	0.3376	0.1533
SVM	0.1183	0.1109
BAYES	0.3521	0.1284

Table 5: Root Mean Square Error Validation for Ortho Diagnosis Data2

Models	Classifier (Single)	Stacked Ensemble
LR	0.3477	0.1162
KNN	0.1466	0.1160
DTC	0.3435	0.1516
SVM	0.1183	0.1095
BAYES	0.3521	0.1284

4.2. Statistical Significance Test

The statistical tests were run at a significance level of 5%, and the results showed that the outcomes of the suggested ensemble stacking with the min-max technique are statistically significant. This article uses a two-tailed paired statistical hypothesis test that is more commonly known as the t-test. This means paired sample t-test, also known as the dependent sample t-test, is a statistical process that can be used to evaluate whether or not there is a difference in mean value between two different sets of observations.

Hypothesis Ho: Predicting both the data diagnosis-1 & 2 using the ensemble stacking method with various machine learning models like Gaussian NB, logistic regression, decision tree regressor, K-Nearest Neighbor algorithm (KNN), and Support Vector Machines (SVM) does not give more accuracy.

Alternate Hypothesis H1: Predicting the data diagnosis-1 & 2 using the ensemble stacking method with various machine learning models like Gaussian NB, logistic regression, decision tree regressor, K-Nearest Neighbor algorithm (KNN), and Support Vector Machines (SVM), using the min-max algorithm gives more accuracy than the ensemble stacking method with various machine learning models.

For this different train, test sets of the same dataset are utilized. Table 2 is used to present the p-value calculated for models based on the results of paired t-test statistical methods. To statistically examine the performance of the ensembled min-max algorithm approach with the categorization of other relationships, a paired sample t-test is used, in which each model is measured twice, resulting in pairs of observations. This provides significant evidence against the null hypothesis (Ho), which states that the proportion of interpreted positively correlated performance metrics created by the ensembled stacking with the min-max algorithm method. If these ratios differ considerably from one another, which would be consistent with the alternative hypothesis, then the ensemble stacking with the minimax algorithm does have a significant impact.

Table 2: p-value for proposed stacking model using t-test

S. No	Stacking Model	Proposed Weighted Min-Max	p-value
Data Diagnosis-1	85.484	97.45	0.0001
	87.097	97.25	
	88.71	97.85	
	82.258	97.9	
	85.484	97.7	
Data Diagnosis -2	82.258	97.3	0.0004
	83.871	97.25	
	87.097	97.05	
	85.484	97.95	
	88.71	97.75	

5. Conclusion

Since the information age, medical data has become increasingly complex, and data mining and medical collaboration are necessary to construct a supplementary diagnostic platform that meets medical data and diagnostic standards for precision and speed. Mining medical data rules yielded an additional diagnostic model for orthopedic clinical data. This research evaluates the ensemble stacking method compared to the random forest approach and the related classification algorithm. And also, suggests a weighted ensemble stacking with Gaussian NB, logistic regression, decision tree regressor, K-Nearest Neighbor algorithm (KNN), and Support Vector Machines (SVM) using a min-max algorithm for developing a prediction model for orthopedic auxiliary classification. The weighted ensemble stacking min-max model outperforms its competitors in terms of processing speed, and it also achieves 97.800% and 97.85% accuracy when applied to the orthopedic diagnosis of disc hernia, normal, and spondylolisthesis. With the peculiarities of medical data and medical diagnosis in mind, it is clear that the weighted ensemble stacking algorithm can handle medical data efficiently and fast. The disease features of the patients are assessed using the qualities of the diagnosis and treatment data, and simple predictions are created to aid the hospital in making successful decisions and the high-incidence population in preventing issues in advance. A well-executed version of this platform has the potential to boost the diagnostic efficiency of orthopedic surgeons, enhance the quality of medical services, advance the reform of medical information, and pave the way for the future growth of smart medical care.

Limitations

The suggested weighted ensemble stacking method for orthopedic diagnosis has disadvantages such as dependence on data quality, potential algorithm bias, decreased interpretability, and the requirement for generalizability testing across various patient populations. Important ethical and regulatory issues include patient privacy and adherence to healthcare standards. Human knowledge integration is crucial, and thorough clinical validation is required to guarantee the dependability and safety of models in real-world healthcare environments. It is essential to overcome these restrictions for the effective implementation and acceptance of the orthopedic diagnostic approach.

Acknowledgments
This research work was supported by DST-RUSA through University of Madras under the RUSA 2.0 project.

References

[1] A. C. Kaidi, J. E. Hellwinkel, M. P. Rosenwasser, W. M. Ricci, The history of orthopaedic surgery in india: from antiquity to present, International Orthopaedics 45 (10) (2021) 2741–2749.
[2] A. K. Jain, Current state of orthopedic education in india, Indian journal of orthopaedics 50 (4) (2016) 341.
[3] B. Murlimanju, P. Krishnaprasad, P. S. Rai, K. Dinesh, L. V. Prabhu, Current state of orthopedic education in india, Indian Journal of Orthopaedics 51 (3) (2017) 349.
[4] S. Li, X. Zhang, Research on orthopedic auxiliary classification and prediction model based on xgboost algorithm, Neural Computing and Applications 32 (7) (2020) 1971–1979.
[5] A. P. Kurmis, J. R. Ianunzio, Artificial intelligence in orthopedic surgery: evolution, current state and future directions, Arthroplasty 4 (1) (2022) 1–10.
[6] X.-G. Han, W. Tian, Artificial intelligence in orthopedic surgery: current state and future perspective, Chinese medical journal 132 (21) (2019) 2521– 2523.
[7] B. Innocenti, Y. Radyul, E. Bori, The use of artificial intelligence in orthopedics: Applications and limitations of machine learning in diagnosis and prediction, Applied Sciences 12 (21) (2022) 10775.
[8] A. Scala, A. Borrelli, G. Improta, Predictive analysis of lower limb fractures in the orthopedic complex operative unit using artificial intelligence: the case study of aou ruggi, Scientific Reports 12 (1) (2022) 1–11.
[9] T. D. Koreckij, J. S. Fischgrund, Degenerative spondylolisthesis, Clinical Spine Surgery 28 (7) (2015) 236–241.
[10] G. Ren, L. Liu, P. Zhang, Z. Xie, P. Wang, W. Zhang, H. Wang, M. Shen, L. Deng, Y. Tao, X. Li, J. Wang, Y. Wang, X. Wu, Machine learning predicts recurrent lumbar disc herniation following percutaneous endoscopic lumbar discectomy, Global Spine Journal (2022) 219256822210976 doi: 10.1177/21925682221097650.
[11] O. K. Oyedotun, E. O. Olaniyi, A. Khashman, Disk hernia and spondylolisthesis diagnosis using biomechanical features and neural network, Technology and Health Care 24 (2) (2016) 267–279.
[12] H. H. Patel, P. Prajapati, Study and analysis of decision tree based classification algorithms, International Journal of Computer Sciences and Engineering 6 (10) (2018) 74–78.
[13] V. Jakkula, Tutorial on support vector machine (svm), School of EECS, Washington State University 37 (2.5) (2006) 3.
[14] M. Awad, R. Khanna, Support vector machines for classification, in: Efficient learning machines, Springer, 2015, pp. 39–66.
[15] H.-A. Park, An introduction to logistic regression: from basic concepts to interpretation with particular attention to nursing domain, Journal of Korean Academy of Nursing 43 (2) (2013) 154–164.
[16] S. R. Kessy, M. Sherris, A. M. Villegas, J. Ziveyi, Mortality forecasting using stacked regression ensembles, Scandinavian Actuarial Journal 2022 (2021) 591 – 626.
[17] K. T. Ngo, Stacking ensemble for auto ml, Ph.D. thesis, Virginia Tech (2018).