Indonesian Journal of Electrical Engineering and Computer Science
Vol. 33, No. 2, February 2024, pp. 1302∼1314
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v33.i2.pp1302-1314

Implementing decision support tool for low-back pain diagnosis and prediction based on the range of motions

Ishaya Gambo[1], Chidozie Mbada[2], Segun Aina[1], Michael Magreola[3], Christopher Agbonkhese[4], Rhoda Ikono[1], Timilehin Ogundare[5], Olasunkami Alimi[6], Francis Saah[1]
[1]Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria
[2]Department of Health Professions, Faculty of Health and Education, Manchester Metropolitan University, Berkeley, UK
[3]Department of Computer Science, University of Arkansas, Little Rock, USA
[4]Department of Digital and Computational Studies, Bates College, Lewiston, USA
[5]Computer Science and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa
[6]Department of Computer Science, Faculty of Computing, Air Force Institute of Technology, Kaduna, Nigeria

Article Info

Article history:

Received Dec 23, 2021
Revised Dec 5, 2023
Accepted Dec 28, 2023

Keywords:

Classification
Decision support system
Low-back pain
Machine learning
Random forest
Stack ensemble
Support vector machine

ABSTRACT

Low-back pain (LBP) is a complex health problem requiring accurate diagnosis and effective treatment. However, the current decision support system (DSS) for LBP only considers the patient's pain intensity and treatment suitability, which may not lead to optimal outcomes. This paper proposes a novel DSS that combines machine learning (ML) and expert input to classify LBP types and provide more reliable and personalized recommendations. We used an open-source dataset to train and test various ML models, including an ensemble model that combines multiple classifiers. We also performed data analysis and feature extraction to enhance the model's predictive power. We developed a prototype tool to demonstrate the model's performance and usability. Our results show that the ensemble model achieved the highest accuracy of 92.02%, followed by random forest (RF) (91.01%), multilayer perceptron (MP) (91.01%), and support vector machine (SVM) (87.88%). Our findings suggest that ML can help LBP specialists diagnose and treat LBP more effectively by learning from historical data and predicting LBP categories. Our DSS can potentially improve the quality of life for LBP patients and reduce the burden on the healthcare system.

Corresponding Author:

Ishaya Gambo
Department of Computer Science and Engineering, Obafemi Awolowo University
Ile-Ife, Nigeria
Email: ipgambo@oauife.edu.ng

1. INTRODUCTION

Medically, low-back pain (LBP) is a prevalent problem that limits peoples' ordinary ways of life and hinders them from scheduling exercises, tasks and activities [1]. As a symptom, LBP has been identified as the foremost noteworthy social burden disease universally [2]. It is not a disease that cannot be validated by an external standard [3], but a public health challenge [4] that affects people of all ages [5], [6]. For that, LBP is ranked sixth in terms of the overall burden of disability and disability-adjusted life years, with a global point prevalence of 9.4% [7], [8]. LBP affects a large proportion of the population in the world, which is why proper diagnosis and treatment are crucial. With about 60% to 90% of the adult population being at risk of developing LBP in their lives [9], it is pertinently essential to establish clear evidence for managing all types and forms of LBP. However, as the authors in [10], [11] observed, the lack of clear proves for one particular

treatment mediation being prevalent to another justifies the need for proper decision-making. Remarkably, the current clinical practice cannot successfully guarantee the effective management of LBP [10] to determine the LBP types and provide more accurate, effective, and reliable diagnoses and treatments. In this context, the use of decision support systems (DSS) for clinical practice is expected [2], [12]. Worthy of mention, DSS in clinical medicine has gotten significant consideration from researchers and practitioners of information systems [1], [13], [14]. Of specific significance is decision-making during medical diagnosis, which is an essential aspect of patient care and administration by healthcare experts [1], [12]-[16].

Against the background described in the preceding paragraph, existing DSS for LBP predominantly focus on assessing pain intensity and treatment suitability, contributing to enhanced medical decision-making and improved care quality [15]. These systems offer clinical guidance for diagnosing and treating LBP, fostering effective doctor-to-patient collaboration and improving communication skills within health information system settings [17]. However, there is a notable absence of classification systems that stratify LBP types based on range of motion for comprehensive clinical decision-making. Consequently, there is an urgent need for an intuitive, user-friendly model to assist physicians in the effective diagnosis and treatment of LBP.

This paper endeavors to develop a model for classifying LBP types based on range of motions, aiming to furnish physicians with more precise, effective, and reliable diagnoses and treatments. The contribution is twofold: methodologically, the paper presents a framework merging machine learning and expert involvement for LBP diagnosis, evaluated using an open-source range of motion dataset. Practically, a prototype tool is implemented, allowing analysis of range of motion features to predict LBP types, demonstrating the framework's capability and suitability. The research outcome delivers a model for a DSS, facilitating clinical decision-making and treatment recommendations for LBP specialists.

2. BACKGROUND AND MOTIVATION

Research on LBP management and treatment has become increasingly significant due to its global impact and rising occurrence [18]. Despite this, medical practitioners face challenges in effectively diagnosing and stratifying LBP types, leading to difficulties in providing appropriate treatment, often attributed to the time-consuming nature of the process [19]. Patient stratification for LBP treatment can be accomplished through classification systems and clinical prediction rules (CPRs), with CPRs relying on statistical analysis to identify key predictors of outcomes or conditions. However, diagnosing LBP based on range of motion impairment remains challenging, requiring expert knowledge of anatomical and physiological complexities [1]. The goal is to introduce a novel approach supporting statistical analysis and the extraction of valuable patterns from LBP data to enhance decision-making. Implementing a clinical decision support system (CDSS) for LBP, especially using machine learning, aims to address these challenges and improve the quality of care effectively [20]. A well-designed CDSS can reduce diagnostic errors, provide fast results, offer essential information to specialists for clinical leadership, and ensure efficiency without replacing experts in the decision-making process [2]. The objective is to leverage machine learning to develop a CDSS tool for LBP diagnosis and prediction based on range of motions, utilizing a classifier to recognize patients' data and determine LBP types.

The rest of this paper is structured as follows. In section 3, the paper reviews related works and establish the research gaps. Section 4 presents the materials and methods explaining the data collection process, model implementation pipeline, model formulation, and evaluation procedure. In section 5, we present the model simulation and implementation. In section 6 describes the results and provides the discussion, and section 7 presents the threats to validity. Finally, in section 8 presents the conclusions and recommendations for future work.

3. RELATED WORK

The literature discusses existing literature on digital platforms for healthcare, with a specific focus on musculoskeletal (MSK) disorders, especially LBP. The literature review highlights various models and approaches, emphasizing their contributions and limitations. Lin *et al.* [1] developed a DSS for LBP diagnosis, addressing uncertainty management and clinical evaluations. Smith *et al.* [21] implemented a CDSS for managing chronic pain at the primary care level. Tascau *et al.* [22] used an integrated DSS for LBP management, emphasizing the reliability of patient information. Pombo *et al.* [23] explored machine learning (ML) to enhance CDSS for LBP diagnosis. Van-Hooff *et al.* [24] focused on CDSS based on logical proof for spine-related consultations.

Gulbandilar *et al.* [25] used a fuzzy logic algorithm to predict LBP intensity, while Navani and Li [26] designed an ML system for chronic pain risk prediction. Bach *et al.* [27] developed a self-back control system for smartphones but faced limitations. Gaonkar *et al.* [28] employed ML techniques for classifying LBP, and Lima *et al.* [29] studied back muscle activity in chronic LBP patients during functional tasks. The summary critiques each study's strengths and weaknesses, such as the specificity of LBP types addressed, accuracy of predictions, and practical applications.

The literature concludes by asserting the study's unique contribution-proposing a model to facilitate clinical decision support for stratifying LBP based on types. The aim is to provide physicians with more accurate and effective diagnoses, ultimately reducing the global burden of disabilities and deaths caused by LBP. The significance of the study lies in its focus on CDSS in LBP, an area that has received less attention despite existing research on decision support systems in health and medical informatics.

4. MATERIALS AND METHOD

The research employed an experiment-based quantitative approach, drawing on simulation and implementation methods as outlined in references [30], [31]. The study's focus was on constructing a decision support model that identifies distinct LBP types through ML, utilizing data primarily derived from secondary sources related to LBP patients. The comprises various stages: data exploration and analysis, feature extraction, model training and testing, model evaluation, and prediction. The conceptual flow highlights the systematic progression followed in the development of the decision support model for LBP type determination.

4.1. Data exploration and analysis

In this study, a dataset focusing on LBP symptoms and range of motion was obtained from the Kaggle Data-Science Repository [32]. A total of 310 patients' details were collected, consisting of 12 input features related to pelvic and spinal parameters. To enhance the dataset's relevance to LBP, a brainstorming session involving six experts from the Department of Medical Rehabilitation at Obafemi Awolowo University and the Orthopedic Unit of Obafemi Awolowo University Teaching Hospital Complex was conducted. These experts, with a minimum of 8 years of post-graduation experience, played a crucial role in identifying features essential for classifying and stratifying LBP. The goal of this session was to provide insights into the dataset, understand its relation to LBP, and assess its suitability for building a decision support model.

The dataset was structured with 12 input features, but for model processing, six features were considered, while the remaining six were automatically checked by the system. The input features, such as pelvic incidence, lumbar lordosis angle, and spondylolisthesis degree, were discussed and clarified during the expert session. The dataset was categorized into two classes (normal and abnormal), representing different severity levels of LBP. The abnormal class was further subdivided into three subclasses (acute, sub-acute, and chronic). The identified input features were utilized to construct a model using three machine learning algorithms for classification. The labeled feature maps, detailing each feature's names and values, were employed in this process. Overall, the research aimed to leverage the expertise of medical professionals to enhance the dataset's relevance and build an effective decision support model for LBP classification.

4.2. Model implementation pipeline

Figure 1 shows the description of the system operations. For pre-processing, the dataset was transformed to include indicators transformation and dimensionality reduction. The feature construction approach was used to improve the ML algorithms' descriptive accuracy [33]. Next, the feature selection process was performed to reduce the dataset and number of features.

The study focused on two key events related to LBP indicators and class labels (normal, abnormal with subclasses of acute, sub-acute, and chronic). A chi-square test was conducted to assess the relationship between each LBP indicator and the class label, determining their independence or necessity. Indicators found to be independent were discarded, while necessary indicators were retained. Following feature selection, the data was split for training and testing using the SciKit library. The training dataset underwent classification using an Ensemble of machine learning algorithms, including support vector machine (SVM), random forest (RF), and multilayer perceptron (MLP). The classification results were evaluated for accuracy, precision, recall, f-measure, and receiver operating characteristics (ROC) metrics, providing a validated decision support tool for predicting the severity level of LBP. For implementing the decision support tool to determine LBP, the Python programming language in the visual studio software coding environment was used.

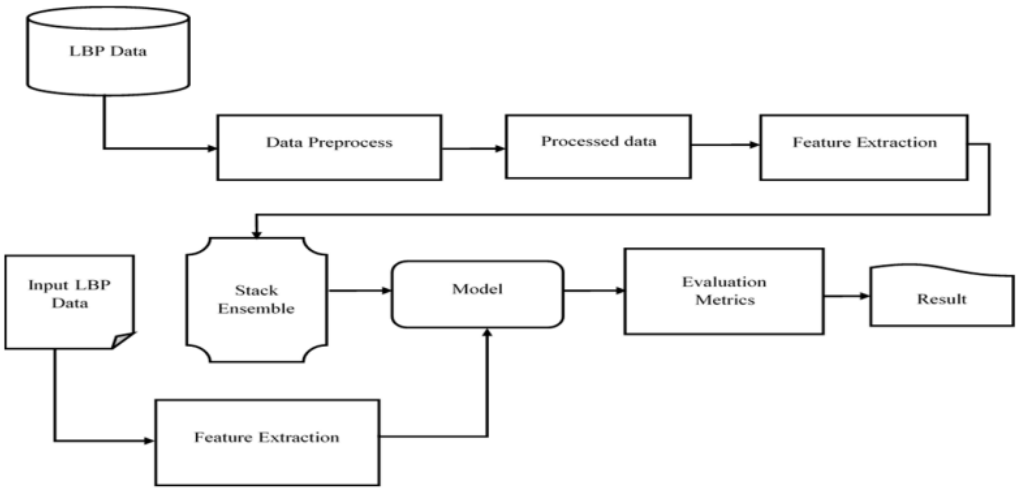


Figure 1. Description of system operation

4.3. Model formulation, simulation and evaluation processes

An ensemble of supervised ML algorithms (SVM, RF, and MLP) was used to formulate the model for classification and implementation of the decision support tool. In particular, the stacked Ensemble that involves training and testing on the dataset based on the classifiers was used. The simulation was carried using the Python programming language. The decision support model was simulated using the training dataset (80% of the dataset) on the stack ensemble method with the scikit-learn simulation tool in the python ML library. For that, the dataset was clustered and classified to determine LBP type. The result of the classified cluster was stored in a cluster.

Additionally, the testing set (20% of the dataset) was applied to the classification model. The model performances were measured using accuracy, precision, recall, f-measure, and ROC. In this context, the accuracy determines the complete correctness of the classifier after prediction, as shown in (1). The precision determines the proportion of the actual class predicted negatively see (2). The recall determines the proportion of the actual cases classified correctly or positively see (3). The f-measure determines the harmonic mean of recall and precision for each class for which the recall and precision were calculated to validate the degree of the test's accuracy see (4). The ROC curve visualizes the classifier's performance to reflect a two-dimensional difference in the test dataset used for testing the decision support model. The two dimensions are the true positive rate (TPR) against the true negative rate (TNR).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TN}{TN + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F - measure = \frac{2 * precision * recall}{precision + recall} \tag{4}$$

As (1), (2), and (3) show, TP is the number of true positives in the dataset. The number of false positives is denoted as FP, TN is the number of true negatives, and FN is the number of false negatives. Worthy of mention, two sets of the labels are built when using precision and recall for the metrics to validate the model. This includes the reference set containing the correct value, and the test set contains values guessed by the classifiers. These two values are compared for each label to determine precision and recall. Overall, a confusion matrix was used to summarize the model validation results of the classification, primarily to determine the TP, TN, FP, and FN values.

5. MODEL SIMULATION AND IMPLEMENTATION

We implemented a tool that facilitates the LBP DSS based on the range of motion test. Making reference to the tool, the range of motion test was performed using the features extracted from the dataset with the LBP DSS. The tool accepts the different observations from the features which the system extracts automatically. The tool predicts the inputted features to determine the severity level of the LBP. The range of motion test was carried out subsequently to determine if the LBP is normal, acute, sub-acute, or chronic using the tool. Table 1 described in summary the range of motion test values the system extracted from the dataset for the features that predicted normal and abnormal (acute, sub-acute, and chronic), respectively. Figures 2-5 show the prediction results reflecting the range of motion test values entered in each field and the severity levels. Figure 2 showed that the predicted result is normal, indicating that given the range of motion test values of the features extracted from the data set, the patient has normal LBP. Figure 3 showed that the predicted result is acute, indicating that given the range of motion test values of the features extracted from the data set, the patient has acute LBP. Figure 4 showed that the predicted result is sub-acute, indicating that the patient has sub-acute LBP given the range of motion test values of the features extracted from the data set. Figure 5 showed that the predicted result is chronic, indicating that given the range of motion test values of the features extracted from the data set, the patient has chronic LBP.

Table 1. Summary of range of motion test values for predicting LBP cases

No	Features	Normal	Abnormal		
			Acute	Sub-acute	Chronic
1	Pelvic incidence (PI)/Slope	9.4251	72.4250	8.0214	8.4574
2	Pelvic tilt (PT)/Direct tilt	19.2541	53.3512	37.8332	18.2456
3	Lumbar lordosis angle (LLA)/Thorocic slope	41.6524	57.2650	79.2542	43.1245
4	Sacral slope (SS)/Cervical tilt	23.2001	56.7412	48.5763	21.2542
5	Pelvic radius (PR)/Sacrum angle	112.2010	130.2451	101.5857	111.2360
6	Spondylolisthesis (S)/Scoliosis slope	10.3654	10.2351	48.45.75	12.2154

Figure 2. Predicted result determining the severity level of LBP as normal

6. RESULTS AND DISCUSSION

The confusion matrix in Figures 6 and 7 contains information about the actual and predicted classifications used to measure the model performance [34], [35]. Figure 6 shows the confusion matrix for the first phase of the dataset used for training (80%). Figure 7 shows the confusion matrix for the second phase of

Figure 3. Predicted result determining the severity level of LBP as acute

Figure 4. Predicted result determining the severity level of LBP as sub-acute

Figure 5. Predicted result determining the severity level of LBP as chronic

the dataset used for testing (20%). We use the confusion matrix to visualize the tasks performed in the classification [36]. As Figures 6 and 7 reflect, the row contains the predicted classes, while the column contains

the model's actual class. Moreover, the predicted values that are correct are labelled diagonally in each of the confusion matrices. With that, it was easy to interpret errors in the prediction obtained from values outside the diagonal visually [36]. Therefore, we used precision, recall, f-measure, and accuracy to evaluate the model's performance for the data shown in Figures 6 and 7. Figure 6 shows that 0–abnormal class and 1-normal. Out of 40 observations dedicated for evaluation, 22 for abnormal, and 18 were normal. The system predicted rightly 21 and misclassified 1 as normal. The system predicted all the 18-observations for normal. All these test sides were not known to the system. These were separated test sets to evaluate the system.

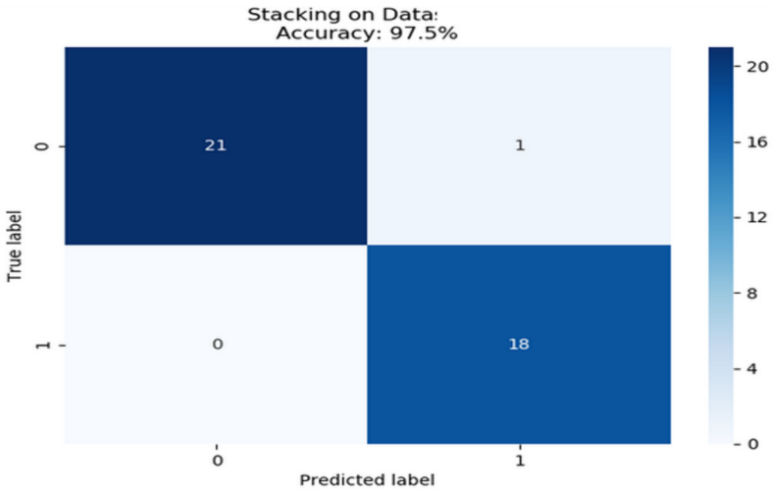


Figure 6. Confusion matrix on the first phase of the dataset used for training (80%)

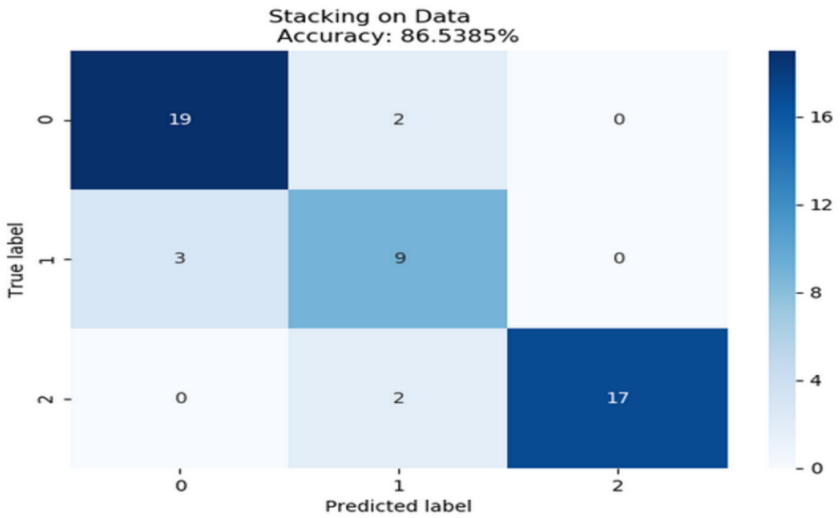


Figure 7. Confusion matrix on the second phase of the dataset used for testing (20%)

Figure 7 indicates the total of 52 test sets that were separated from the dataset. There are 3 classes available (0-acute, 1-chronic, and 2-subacute). The system classified 19 right out of 21 observations available for acute and misclassified 2 as chronic. 12-observations or records were available for chronic. The system predicted 9 right and misclassified 3 as acute. Out of 19 observations available for sub-acute, the system predicted 17 observations right and misclassified 2 as chronic. Table 2 shows the summary of the classification results of LBP.

Table 2. Classification results of LBP

Classes	0	**1**	2	Total observation
Acute	19	2	0	21
Chronic	3	9	0	12
Sub-acute	0	2	17	19

Going further, Table 3 shows the results based on the first phase of the dataset used. At the same time, Table 4 display the result based on the second phase of the dataset used. Tables 5 and 6 show the results of each ML algorithms (SVM, RF, and MLP) and the stacked ensemble.

Table 3. Evaluation metrics using training set

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
MLP	95.00	95.00	95.46	95.99
SVM	97.50	97.34	97.73	97.49
RF	97.50	97.83	97.22	97.46
Stacking	97.50	97.39	97.73	97.73

Table 4. Evaluation metrics using test set

Classifier	Accuracy (%)	Precision (%)	Recall (%)	F-measure (%)
MLP	80.77	81.51	82.60	80.49
SVM	84.62	84.09	85.78	84.70
RF	86.54	84.74	97.22	84.98
Stacking	86.54	83.09	82.21	82.21

Table 5. Differences in ensemble model evaluation over training and test phases

Metrics	Phase 1 (%)	Phrase 2 (%)	Differences (%)
Accuracy	97.50	86.54	10.96
Precision	97.39	83.09	14.30
Recall	97.73	82.21	15.52
F-measure	97.73	82.21	15.52

Table 6. Model evaluation result

	Metrics	Model performance
1	Accuracy	92%
2	Precision	80%
3	Recall	90%
4	F-measure	85%

Tables Tables 3 and 4 show how sensitive the system is to precision. However, there have been different results with minor changes based on the random shuffling of the datasets at different program execution. As Tables 5 and 6 reflect, RF performed relatively better than the other algorithms. For the RF, the hidden-layer-size specified was (500x500), which implies 500 hidden units with two layers is for RF. The number of times and the hidden layers and units specified can improve the performance of the system. Nevertheless, it is essential that we avoid overfitting the model. The result in the stacking drop down is the performance of the other algorithm. We used the ROC curve to visualize and inspect the performance of the algorithms. The ROC curve compares the rate at which the classifier is making a correct prediction, "true positive" (TP) prediction and the "false positive" (FP) prediction. Figure 8 shows that the system had a good prediction because the system had a perfect trade-off between TPR and FPR. The prediction is approximately 1, which is above the line. Figure 9 shows that all the classes as predicted by the stacked Ensemble were above the line. Since the system had three classes for these data, it has three curves above the line, and the area under curve (AUC) is shown on the right-hand side. The AUC is an aggregation metric that determines how well the system prediction was made in this context. Figure 10 shows the cross-validation result, which is the other way of detecting the performance of the dataset. Cross-validation helps the system know the model's performance in real-time, not just for testing and splitting data.

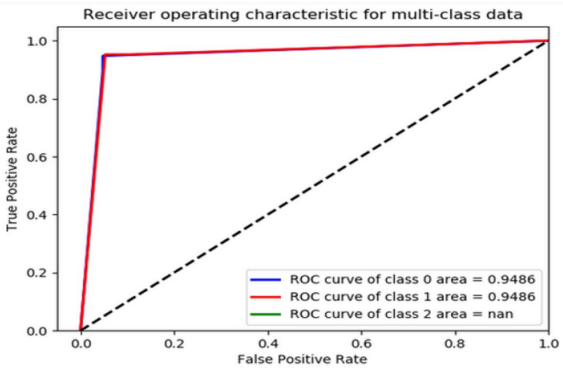


Figure 8. ROC for the ensemble model on the first phase of the data

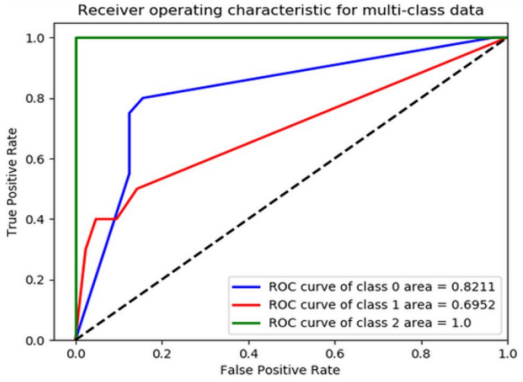


Figure 9. ROC for the ensemble model on the second phase of the data

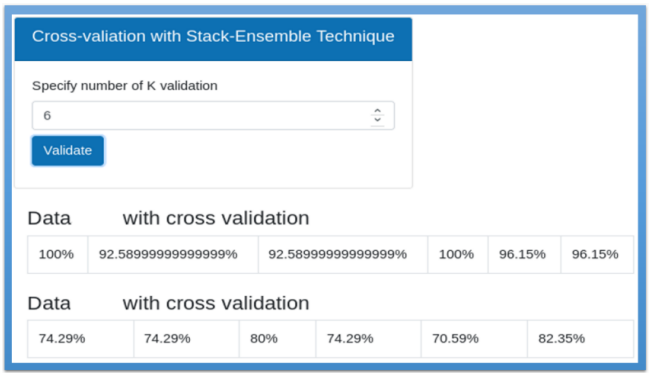


Figure 10. Cross-validation on both phases of the dataset

Table 5 presents the data's evaluation result for both phases and their percentage difference. Table 5 shows that the model has a higher detection accuracy of 97.5% based on the phases of data, with 84.615% accuracy. This shows that the first phase of data has a difference of 13.885% higher than the second phase of data. The precision metrics for the model, as shown in Table 5, indicated that the first dataset has 97.68%. In comparison, the data has 83.946% precision in the second phase. This means that the first phase of the data has a difference of 13.422%. The recall evaluation metrics showed that the first phase of the data had a higher recall of 97.727% than the second phase of data with 83.229%. This means that the first phase of the data has a difference of 14.043%. F-measure showed that the first phase of the data had a higher F-measure of 97.727% compared to the second phase of data with 83.229%. This means that the first phase of the data had a difference of 14.043%. The developed model accumulated an accuracy of 92.03%; for precision, the model has a precision of 80%. The recall evaluation metrics showed that the model has a recall of 90.94%, and for F-measure, the model has a rate of 85%. Table 6 shows detail of the ensemble model performance.

As Table 5 shows, the differences in terms of the evaluation metrics between both phases of the data are relatively low. This means that the model is good enough for prediction and classification purposes. In particular, with an accuracy difference of 10.96%, the experts accept that the results were good enough, as there is no perfect system. Also, the experts showed much confidence and trust in using the tool, especially for training and educational purposes in the laboratories, to show how the severity level of LBP can be determined. SVM and RF produce the same accuracy of 91.01%, respectively. Likewise, the ensemble model outperformed SVM, MLP, and RF with 92.02% accuracy, as seen in Table 7. The stacked ensemble model achieved 92.02% accuracy. Consequently, it means the model can correctly predict and determine the severity level of LBP based on the range of motion test performed.

Table 7. Overall accuracy for machine learning models

	Model	Accuracy
1	Accuracy	87.89%
2	Precision	91.01%
3	Recall	91.01%
4	F-Measure	92.02%

7. THREATS TO VALIDITY

The validity of the study faces two main threats: internal and external. Internally, the risk lies in collecting, analyzing, and understanding the dataset and its features. This was addressed by involving experienced Orthopedic consultants and specialist physiotherapists in LBP during the study, ensuring a comprehensive understanding and guidance on the necessary features for LBP classification. Additionally, the internal validity threat related to result acceptance and confidence was mitigated by demonstrating the implemented tool to the experts, showcasing the technological and scientific processes involved. External validity, the second threat, underscores the need for validation with real-life datasets in the problem domain and adherence to software design principles for future LBP prediction and classification systems, as emphasized in [37], [38].

8. CONCLUSION

This interdisciplinary study integrates health research with computer science to develop a ML-based DSS for classifying LBP types. The primary goal is to improve the precision, efficiency, and reliability of LBP diagnoses and treatments by identifying key features governing LBP types. The study reviews existing literature on ML-based DSS for LBP, emphasizing their efficacy in detection and categorization. Addressing gaps in prior research, the study conducts feature analysis to extract crucial features for LBP categorization. Utilizing a modified Delphi approach with experts from Obafemi Awolowo University, datasets are analyzed. ML models, particularly an ensemble model, demonstrate superior performance based on accuracy, precision, recall, and F-measure metrics. The developed DSS proves promising for real-world applications, guiding patient treatment and classifying LBP into normal, acute, subacute, or chronic categories, contributing to the evolving field of ML in healthcare. For future work, we recommend the following. Increasing the model parameters and the number of datasets with more features can enhance the model's performance. However, this will cause getting data from more sources and increase processing capacity. Increasing the model's parameter set would indicate a longer time of model training with additional information, ultimately enhancing model performance. We investigate how mutual knowledge can code the relationship between input and output variables. As a result, employing a probabilistic graphical model architecture, such as a Bayesian network to estimate the probability distribution between the input and target variables will assist researchers in identifying the direct influence of each input variable on the target variable. Extend the LBP DSS predictions to include health-related implications of an LBP type. By so doing, more information is offered to the physician, allowing them to deliver more competent advice to patients.

ACKNOWLEDGMENT

The authors would like to thank the Africa Centre of Excellence OAK-Park for the support in executing the research.

REFERENCES

[1] L. Lin, P. J. H. Hu, and O. R. Liu Sheng, "A decision support system for lower back pain diagnosis: uncertainty management and clinical evaluations," *Decision Support Systems*, vol. 42, no. 2, pp. 1152–1169, 2006, doi: 10.1016/j.dss.2005.10.007.

[2] V. M. H. Coupé, M. L. van Hooff, M. de Kleuver, E. W. Steyerberg, and R. W. J. G. Ostelo, "Decision support tools in low back pain," *Best Practice and Research: Clinical Rheumatology*, vol. 30, no. 6, pp. 1084–1097, 2016, doi: 10.1016/j.berh.2017.07.002.

[3] L. Manchikanti, "Epidemiology of low back pain," *Pain physician*, vol. 3, no. 2, pp. 167–192, 2000.

[4] C. M. Williams *et al.*, "Low back pain and best practice care a survey of general practice physicians," *Archives of Internal Medicine*, vol. 170, no. 3, pp. 271–277, 2010, doi: 10.1001/archinternmed.2009.507.

[5] C. E. Mbada *et al.*, "Comparative efficacy of clinic-based and telerehabilitation application of mckenzie therapy in chronic low-back pain," *International Journal of Telerehabilitation*, vol. 11, no. 1, pp. 41–58, 2019, doi: 10.5195/ijt.2019.6260.

[6] L. Manchikanti, V. Singh, F. J. E. Falco, R. M. Benyamin, and J. A. Hirsch, "Epidemiology of low back pain in adults," *Neuromodulation*, vol. 17, no. S2, pp. 3–10, 2014, doi: 10.1111/ner.12018.

[7] D. Hoy et al., "The global burden of low back pain: estimates from the Global burden of disease 2010 study," *Annals of the Rheumatic Diseases*, vol. 73, no. 6, pp. 968–974, 2014, doi: 10.1136/annrheumdis-2013-204428.

[8] C. E. Mbada *et al.*, "Comparative effects of clinic and virtual reality-based McKenzie extension therapy in chronic non-specific low-back pain," *Human Movement*, vol. 20, no. 3, pp. 66–79, 2019, doi: 10.5114/hm.2019.83998.

[9] J. K. Arnold Wong and D. Samartzis, "Low back pain in older adults: risk factors, management options and future directions," in *National Library of Medicine*, 2017, vol. 12, doi: 10.1186/13013-017-0121-3.

[10] N. Meziat Filho, "Changing beliefs for changing movement and pain: classification-based cognitive functional therapy (CB-CFT) for chronic non-specific low back pain," *Manual Therapy*, vol. 21, pp. 303–306, 2016, doi: 10.1016/j.math.2015.04.013.

[11] J. B. Staal, R. De Bie, H. C. W. De Vet, J. Hildebrandt, and P. Nelemans, "Injection therapy for subacute and chronic low-back pain," *Cochrane Database of Systematic Reviews*, vol. (3), no. 3, p. CD001824, 2008, doi: 10.1002/14651858.CD001824.pub3.

[12] S. Jansen-Kosterink, L. van Velsen, and M. Cabrita, "Clinician acceptance of complex clinical decision support systems for treatment allocation of patients with chronic low back pain," *BMC medical informatics and decision making*, vol. 21, no. 1, p. 137, 2021, doi: 10.1186/s12911-021-01502-0.

[13] M. W. M. Jaspers, M. Smeulers, H. Vermeulen, and L. W. Peute, "Effects of clinical decision-support systems on practitioner performance and patient outcomes: a synthesis of high-quality systematic review findings," *Journal of the American Medical Informatics Association*, vol. 18, no. 3, pp. 327–334, 2011, doi: 10.1136/amiajnl-2011-000094.

[14] W. Wiharto, Clinical decision support systems theory and practice, vol. 7, no. 2. Springer, 2018.

[15] R. J. Conejar and H.-K. Kim, "A medical decision support system (DSS) for ubiquitous healthcare diagnosis system," in *International Journal of Software Engineering and Its Applications*, 2014, vol. 8, no. 10, pp. 237–244, [Online]. Available: http://dx.doi.org/10.14257/ijseia.2014.8.10.22.

[16] R. Maheronnaghsh, S. Nezareh, M. K. Sayyah, and V. Rahimi-Movaghar, "Developing SNOMED-CT for decision making and data gathering: a software prototype for low back pain," *Acta Medica Iranica*, vol. 51, no. 8, pp. 548–553, 2013.

[17] I. P. Gambo and A. H. Soriyan, "ICT implementation in the Nigerian healthcare system," *IT Professional*, vol. 19, no. 2, pp. 12–15, 2017, doi: 10.1109/MITP.2017.21.

[18] L. D. Morris, K. J. Daniels, B. Ganguli, and Q. A. Louw, "An update on the prevalence of low back pain in Africa: a systematic review and meta-analyses," *BMC Musculoskeletal Disorders*, vol. 19, no. 1, 2018, doi: 10.1186/s12891-018-2075-x.

[19] S. E. Forseen and A. S. Corey, "Clinical decision support and acute low back pain: evidence-based order sets," *Journal of the American College of Radiology*, vol. 9, no. 10, pp. 704-712.e4, 2012, doi: 10.1016/j.jacr.2012.02.014.

[20] A. Wright et al., "Development and evaluation of a comprehensive clinical decision support taxonomy: comparison of front-end tools in commercial and internally developed electronic health record systems," *Journal of the American Medical Informatics Association*, vol. 18, no. 3, pp. 232–242, 2011, doi: 10.1136/amiajnl-2011-000113.

[21] M. Y. Smith, J. D. Depue, and C. Rini, "Computerized decision-support systems for chronic pain management in primary care," *Pain Medicine*, vol. 8, no. SUPPL.3, pp. S155–S166, 2007, doi: 10.1111/j.1526-4637.2007.00278.x.

[22] M. Toth-Tascau, D. I. Stoia, and D. Andrei, "Integrated methodology for a future expert system used in low back pain management," in *SACI 2012 - 7th IEEE International Symposium on Applied Computational Intelligence and Informatics, Proceedings*, 2012, pp. 315–320, doi: 10.1109/saci.2012.6250022.

[23] N. Pombo, P. Araújo, and J. Viana, "Knowledge discovery in clinical decision support systems for pain management: a systematic review," *Artificial Intelligence in Medicine*, vol. 60, no. 1, pp. 1–11, 2014, doi: 10.1016/j.artmed.2013.11.005.

[24] M. L. Van Hooff, J. Van Loon, J. Van Limbeek, and M. De Kleuver, "The Nijmegen decision tool for chronic low back pain. Development of a clinical decision tool for secondary or tertiary spine care specialists," *PLoS ONE*, vol. 9, no. 8, p. e104226, 2014, doi: 10.1371/journal.pone.0104226.

[25] E. Gulbandilar, M. Sari, and A. Cimbiz, "Prediction of low back pain using a fuzzy logic algorithm," no. September, pp. 1–6, 2015, doi: 10.13140/RG.2.1.3929.6486.

[26] A. Navani and G. Li, "Chronic pain challenge: a statistical machine-learning method for chronic pain assessment," *Journal on Recent Advances in Pain*, vol. 2, no. 3, pp. 82–86, 2016, doi: 10.5005/jp-journals-10046-0048.

[27] K. Bach, T. Szczepanski, A. Aamodt, O. E. Gundersen, and P. J. Mork, "Case representation and similarity assessment in the SELFBACK decision support system," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016, vol. 9969 LNAI, pp. 32–46, doi: 10.1007/978-3-319-47096-2_3.

[28] A. P. Gaonkar, R. Kulkarni, R. D. Caytiles, and N. C. S. Iyengar, "Classification of lower back pain disorder using multiple machine learning techniques and identifying degree of importance of each parameter," *International Journal of Advanced Science and Technology*, vol. 105, pp. 11–24, 2017, doi: 10.14257/ijast.2017.105.02.

[29] M. Lima, A. S. Ferreira, F. J. J. Reis, V. Paes, and N. Meziat-Filho, "Chronic low back pain and back muscle activity during functional tasks," *Gait and Posture*, vol. 61, pp. 250–256, 2018, doi: 10.1016/j.gaitpost.2018.01.021.

[30] C. I. Chang and H. Ren, "An experiment-based quantitative and comparative analysis of target detection and image classification algorithms for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 2 II, pp. 1044–1063, 2000, doi: 10.1109/36.841984.

[31] V. Arghode, "Qualitative and quantitative research: paradigmatic differences," *Global Education Journal*, vol. 2012, no. 4, pp. 155–163, 2012, doi: http://www.franklinpublishing.net/globaleducation.html.
[32] M. John peloza, "Lower back pain symptoms," *SPINE-health*, 2017. https://www.kaggle.com/sammy123/
[33] R. Alfred, "Feature transformation: a genetic-based feature construction method for data summarization," *Computational Intelligence*, vol. 26, no. 3, pp. 337–357, 2010, doi: 10.1111/j.1467-8640.2010.00362.x.
[34] I. Gambo and K. Taveter, "Identifying and resolving conflicts in requirements by stakeholders: a clustering approach," in *International Conference on Evaluation of Novel Approaches to Software Engineering, ENASE - Proceedings*, 2021, vol. 2021-April, pp. 158–169, doi: 10.5220/0010526901580169.
[35] I. Gambo and K. Taveter, "Stakeholder-centric clustering methods for conflict resolution in the requirements engineering process," in *Communications in Computer and Information Science*, vol. 1556 CCIS, 2022, pp. 183–210.
[36] I. Gambo, A. Adjicheboutou, R. Ikono, I. Olaronke, and T. Yange, "An investigative process model for predicting information diffusion on social media: information system perspective," *Ife Journal of Technology*, vol. 27, pp. 47–59, 2020.
[37] E. Halilaj, A. Rajagopal, M. Fiterau, J. L. Hicks, T. J. Hastie, and S. L. Delp, "Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities," *Journal of Biomechanics*, vol. 81, pp. 1–11, 2018, doi: 10.1016/j.jbiomech.2018.09.009.
[38] I. Gambo, R. Massenon, T. S. Yange, R. Ikono, T. Omodunbi, and K. Babatope, "Software design specification and analysis of insulin dose to adaptive carbohydrate algorithm for type 1 diabetic patients," in *Studies in Computational Intelligence*, vol. 932, Springer, 2021, pp. 107–132.

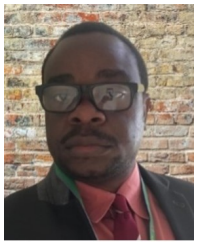
BIOGRAPHIES OF AUTHORS

Ishaya Gambo research is in the area of software engineering (SE), particularly in requirements engineering (RE), software testing and software architecture. He appreciates applying his research in the healthcare domain. The emphasis of his research is on user's and developers' perspectives of software systems. Ishaya has experience of applied projects and has a clear history of the full research life cycle, including academic publishing in journals and extensive presentation at conferences. He can be contacted at email: ipgambo@oauife.edu.ng.

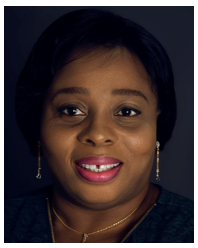
Chidozie Mbada a Senior Lecturer in the Department of Health Professions, Manchester Metropolitan University, United Kingdom. He is into digital applications of mobile technology, virtual reality, and machine learning/artificial intelligence into health care research and practice. Specifically, he has made novel contributions in telerehabilitation where alongside his colleagues, have developed self-treatment mobile App for patients with low-back pain, cartoon-based treatment application, as well as a virtual reality game for low-back rehabilitation. He can be contacted at email: c.mbada@mmu.ac.uk.

Segun Aina is a Senior Lecturer in the Department of Computer Science Engineering with specific research interest in digital signal processing, pattern classification, image processing using ML techniques and pervasive and mobile computing. He can be contacted at email: s.aina@oauife.edu.ng.

Michael Magreola is a postgraduate student at the Department of Computer Science, University of Arkansas at Little Rock, 2801 S. University Ave, USA. He has interest in Artificial Intelligence research. He is passionate about the application of AI-based systems to address real-world challenges. He can be contacted at email: mmagreola@ualr.edu.



Christopher Agbonkhese is a Visiting Assistant Professor in the Department of Digital and Computational Studies, Bates College, Lewiston, USA. His research encompasses software systems, health informatics and data analytics, where he employs techniques like machine learning and data mining to construct models for addressing critical challenges. He can be contacted at the email: cagbonkhese@bates.edu.



Rhoda Ikono is a Senior lecturer at the Department of Computer Science and Engineering, Obafemi Awolowo University Ile-Ife. Her research interests are information system, health informatics, biomedical informatics, human robot interaction, interoperability and ontology matching, and software product usability. She has published widely in international journals and conferences in the field of computer science. She can be contacted at email: rudo@oauife.edu.ng.

Timilehin Ogundare research interest spans the filed of Artificial Intelligence, Business Intelligence Analytics, Health Informatics, Data Science and Web programming. He is more passionate about the application of AI-based systems to address real-world challenges. He can be contacted at email: olutimmy93@gmail.com.

Olasunkami Alimi research and development is in the areas of Software Engineering, Software Development, Information Systems, and Algorithms for different areas of human endeavours. He is a member of Nigeria Computer Society (NCS), Computer Professional Registration Council of Nigeria (CPN), Teachers Registration Council of Nigeria (TRCN) and Academia in Information Technology Profession (AITP). He can be contacted at email: alimiom@afit.edu.ng.

Francis Saah completed his Master's degree in Information System at the Department of Computer Science and Engineering, Obafemi Awolowo University, Nigeria. His research interest is in the area of medical and health informatics system. He can be contacted at email: francisssaaah@gmail.com.