JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

GAIT KINEMATICS FOR LOW BACK PAIN: AN AI FRAMEWORK FOR QUANTIFYING PAIN INTENSITY AND PAIN-RELATED DISABILITY

[1]Debarpita Santra, [2]Indraneel Mukhopadhyay
[1]Assistant Professor, [2]Professor
Amity Institute of Information
Technology Amity University
Kolkata, India

Abstract: Despite being one of the most common and incapacitating musculoskeletal disorders in the world, low back pain (LBP) is difficult to diagnose and treat since 85–95% of cases are non-specific in nature. This study uses sagittal plane gait kinematics taken from 2D video recordings to offer a novel machine learning framework for objectively estimating pain severity and pain-related disability in LBP patients. Traditional gait analysis approaches, though effective, are typically resource-intensive and inaccessible in low-resource situations. The proposed method makes use of spatiotemporal and kinematic gait characteristics, such as stride length, cadence, joint angles, asymmetry indices, and step variability that are derived from recorded gait sequences using pose estimation techniques in order to get around these restrictions. A dataset of 100 LBP patients was used to train a Gradient Boosting regression model (XGBoost), which included self-reported pain and disability scores in addition to gait-derived features. The model demonstrated strong predictive performance, with impairment scores predicted with a comparably low error on a 100-point scale and pain intensity estimations on a 10-point scale falling within ±1.0 point of the self-reported values. In line with the established biomechanical consequences of LBP, feature importance analysis revealed that the most informative predictors were cadence, stride length, knee angle, and hip range of motion. The outcomes confirm that AI-enabled gait analysis has the potential to be a scalable and affordable substitute for traditional pain assessment instruments.

Keywords: *Low Back Pain; Gait Kinematics; Pain Intensity; Pain-Related Disability; Gradient Boosting (XGBoost)*

I. INTRODUCTION

Low Back Pain (LBP) is one of the most common and disabling musculoskeletal conditions, with approximately 70% to 85% of individuals globally experiencing it at some point in their lives. [1] Its high prevalence has positioned it as a leading cause of physical impairment, surpassing conditions like depression, diabetes, and osteoarthritis in global disability impact according to the World Health Organization. [1,3] Beyond its clinical toll, LBP imposes a serious economic burden, particularly in industrialized nations. In the United States alone, it accounts for the loss of over 149 million workdays annually and ranks as the third-highest healthcare expenditure at approximately $87.6 billion. [4]

What makes LBP particularly difficult to manage is that 85% to 95% of cases are classified as non-specific, meaning that the exact anatomical source of pain cannot be reliably identified through conventional diagnostic tools. [5,6] This diagnostic uncertainty contributes to varied clinical interpretations, leading to misdiagnoses, delayed interventions, or inappropriate treatment, especially in areas with limited access to imaging or specialist care. [7][8]

Gait alteration is one of the significant signs that is frequently observed in LBP patients. These patients experience pain and discomfort while walking. The compensatory adaptations during walking result in measurable deviations from the typical gait patterns of healthy individuals. [9][10] The deviations can be measured through some important biomechanical variables like stride length, cadence, and joint motion during the stance and swing phases of gait. [11] Hence, gait analysis acts as a reliable, objective, and non-invasive approach for assessing pain intensity and pain related disability, in contrast to the widely used subjective pain intensity and disability assessment tools like Numeric Rating Scale (NRS) [12] and Roland Morris Disability Questionnaire (RMDQ) [13] respectively.

Several studies have investigated the gait patterns in patients with chronic LBP-related conditions. da Fonseca et al., [14] noted that a six-week Pilates rehabilitation program resulted into changes of gait symmetry, pelvic position, and lumbar range of motion, which were assessed by a pre- and post-intervention laboratory gait analysis. In their study, Khodadadeh and Eisenstein, [15] showed that surgical operation increased stride length, joint kinematics, loading symmetry, pain, and functional outcomes among

LBP patients in presence of pre-surgical and post-surgical gait measurements. According to Vickers et al., [16] observer effects were observed when gait was assessed and indicated that the presence of an investigator had an effect on the stride length and the angles of the joints especially among the chronic LBP patients. To make it more accessible, Chan et al., [17] confirmed smartphone-based gait analysis with embedded accelerators and gyroscopes and reported a high correlation with laboratory systems of measuring the joint angles and temporal gait variables. Also, Mullerpatan et al., [18] discovered that there are considerable differences in the stride length and joint movement pattern of Bharatanatyam dancers with and without LBP, which indicates occupational biomechanical effects on LBP.

Though gait analysis plays an important role in assessing non-specific LBP, the analysis methods require a skilled personnel, high-end motion capture systems, and controlled lab environments. But these resources are inaccessible in many healthcare settings, especially in the remote areas or low-resource regions. [19][20] To resolve these issues, this paper proposes a cost-effective and scalable Artificial Intelligence (AI) based solution that accepts 2D sagittal gait video recordings for evaluation of pain severity and associated disability in LBP patients. [21,22]

The AI-based low cost solution involves Gradient Boosting Machine (GBM) algorithms, especially XGBoost, which can efficiently capture the complex, non-linear relationships between the clinical outcomes and the biomechanical features. [23,24] The models are trained with different important spatiotemporal parameters like stride length, cadence, stance-swing ratio, knee and hip joint angles as well as the pain intensity and disability scores which are collected through clinical questionnaires. [22] XGBoost builds an ensemble of decision trees, where each subsequent tree focuses on correcting the residuals of the previous ones, making the model highly interpretable and robust in structured data scenarios. [25,26]

The proposed predictive model was validated with sagittal gait video recording of 100 LBP patients which were taken from the archive of a city based referral government hospital in India with proper ethical consent. With impairment scores predicted with a relatively low error on a 100-point scale and pain intensity estimates on a 10-point scale falling within ±1.0 points of the self-reported values, the model showed high predictive performance. According to feature importance analysis, cadence, stride length, hip range of motion, and knee angle emerged as the most important predictors. The promising results are consistent with the known biomechanical effects of LBP, and suggest that such AI-driven tools could offer real-time, interpretable, and non-invasive support for pain assessment in low-resource clinical environments.

The rest of the paper is organized as follows: Section 2 presents the related work, Section 3 describes the proposed methodology, Section 4 evaluates the performance of the model, and Section 5 concludes the paper.

II. METHODS

A GBM-based predictive framework has been proposed in this paper that utilizes spatiotemporal gait parameters for quantifying pain intensity and disability levels of LBP patients. The gait parameters have been computed using Pose Estimation Algorithm, [21] that was applied on gait video recording of LBP patients. Before training of the model, extraction of the following clinically relevant feature variables from kinematic and spatiotemporal gait parameters, as shown in Table 1, is done.

Table 1: Extracted kinematic and spatiotemporal gait parameters from gait analysis at sagittal plane

Sl No.	**Feature Name**	**Definition**
	Independent Variables	
1	Stride Length (m)	Distance between two consecutive steps of the ipsilateral foot
2	Cadence (steps/min)	Steps per minute
3	Mean Knee Angle (°)	Average of the flexion-extension angle of knee during gait
4	Hip Angle Range (°)	Range of Motion (RoM) of the hip joint
5	Asymmetry Index (%)	Numerically measured imbalance between the left and right gait dynamics
6	Step Time Variability (s)	Standard Deviation of the step duration during different gait cycles
7	Joint Angular Velocity (°/s)	Rate of change of the hip and knee angles as observed in different video frames
	Dependent Variables	
8	Pain Intensity	LBP patient's self-reported pain severity on Numeric Rating Scale (NRS) from 0 to 10
9	Disability Score	LBP patient's disability score between 0 to 100, both inclusive, measured as Oswestry Disability Index (ODI)

Prior to training of the GBM, the dataset undergoes pre-processing. K-nearest neighbour (KNN) algorithm has been used for imputation of missing values. Z-score has been used for outlier detection. Here, no standardization is needed as Gradient Boosting is invariant to feature scaling.

The proposed prediction model is based on Gradient Boosting Regression which was implemented using XGBoost. The reason behind choosing XGBoost Regressor is that it generally performs well on structured tabular data and also is capable to handle missing data, non-linear relationships or multi-collinearity present in the data. The model hyperparameters have been configured as follows:

n_estimators: 100, implying that 100 trees have been used in the boosting sequence
max_depth: 4, implying that the maximum depth of every regression tree has been restricted to 4 in order to manage model complexity
learning_rate: 0.1, implying that this moderate learning rate would maintain the balance between the overfitting risk and the convergence speed
random_state: 42, implying that the results would be reproduced

Gradient Boosting is a sequential ensemble of M (> 0) decision trees, with every tree attempting to minimize the residual error from previous iteration. During training, the iterative residual minimization was done by XGBoost.

Formally, the training dataset is represented as {(xi, yi)}i=1 to N. With the given training data, the additive function is learnt by the model as shown in Eq. (i).

$$Fm(X) = \Sigma m=1 \text{ to } M\ hm(X) \quad \dots (i)$$

Here, Fm(X) represents the prediction following m ($1 \leq m \leq M$) boosting rounds, and hm(X) is the m-th regression tree that is trained on the residuals of the previous ensemble. The model is updated as per the equation shown in Eq. (ii).

$$Fm(X) = Fm\text{-}1(X) + \eta \cdot hm(x) \quad \dots (ii)$$

Here, η represents the learning rate. The iterative additive approach makes the model capable to improve the predictive accuracy at every stage without rapid overfitting.

The model training and testing dataset ratio is 80:20. After training of the model, it is tested and validated to find the accuracy and clinical usefulness. For validation of the proposed predictive model, sagittal gait video recording of 100 LBP patients were taken from the archive of a city based referral government hospital in India with proper ethical consent. The independent and dependent variables as mentioned in Table 1 have been extracted from the video recordings. Table 2 shows a subset of 10 patient records that include two demographic parameters namely age and sex, the gait kinematic as well as spatiotemporal parameters, and the self-reported pain intensity along with calculated disability score. Patient anonymity has been maintained throughout the study. Male LBP patient has been denoted as 0 and female LBP patient has been denoted as 1.

Table 2: Subset of 10 LBP patients' records involving demographic parameters, gait kinematic parameters, and self-reported pain intensity and disability score

Age	Sex	Stride Length	Cadence	Stance-Swing Ratio	Mean Knee Angle (°)	Hip Angle Range (°)	Asymmetry Index (%)	Step Time Variability (s)	Joint Angular Velocity (°/s)	Pain Intensity (NRS)	Disability Score
64	1	1.31	101.48	1.76	152.52	23.54	-0.62	0.194	169.08	9	57.73
71	0	1.21	95.66	1.75	160.26	31.33	11.23	0.227	115.98	9	70.73
42	0	1.29	96.61	1.83	152.43	32.26	12.23	0.184	118.06	8	70.07
49	0	1.13	105.19	1.72	148.05	34.16	4.50	0.281	119.60	9	72.39
28	0	1.09	99.52	1.72	156.98	32.83	3.83	0.160	125.12	9	71.83
31	0	1.25	105.96	1.54	148.66	18.61	16.51	0.217	136.72	8	95.24
26	0	1.16	104.50	1.39	146.53	36.72	-2.86	0.268	147.14	9	58.33
32	0	1.33	119.34	1.32	152.34	27.79	16.75	0.227	121.93	7	68.27
35	1	1.09	93.91	1.68	153.90	32.36	8.24	0.207	147.10	10	87.46
63	0	1.12	99.88	1.95	143.37	30.89	16.07	0.203	113.84	10	100.00

As shown in Table 2, the Pain Intensity on NRS ranges from 7 to 9, which reflects moderate to severe pain especially for chronic LBP patients. The shorter Stride Length and lower Cadence values are found to be associated with higher pain intensity and pain-related disability scores.

III. RESULTS

For evaluating performance of the model, it has been observed how the trained machine learning model predicts pain intensity and pain-related disability. Figure 1.1 shows the graph to reflect Actual vs Predicted Pain Intensity, Figure 1.2 shows the graph for Residuals vs Predicted Pain Intensity, and Figure 1.3 shows the graph for feature importance for prediction of pain intensity.

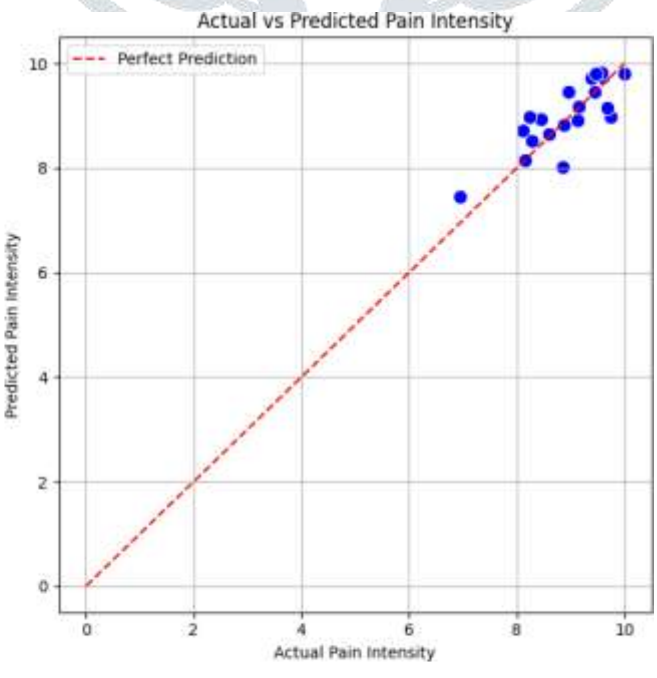

Fig. 1.1. Graph showing Actual vs Predicted Pain Intensity of the model

In Figure 1.1, self-reported pain intensity has been shown along x-axis, and the predicted pain intensity has been shown along the y-axis. Each blue point corresponds to one LBP patient in the test dataset. Any point over the red diagonal line indicates that the prediction is exactly same as the actual pain intensity value. Figure 1.1 shows that there is tight clustering around the diagonal red line, implying that the prediction of pain intensity values are highly accurate. The presence of small deviations between the actual and the predicted pain intensity values suggest model generalization and avoidance of large prediction errors. Low spreading of blue points correspond to low MAE and low RMSE. Here, no outliers are present. As the model shows high accuracy in pain intensity prediction, it can be clinically useful.

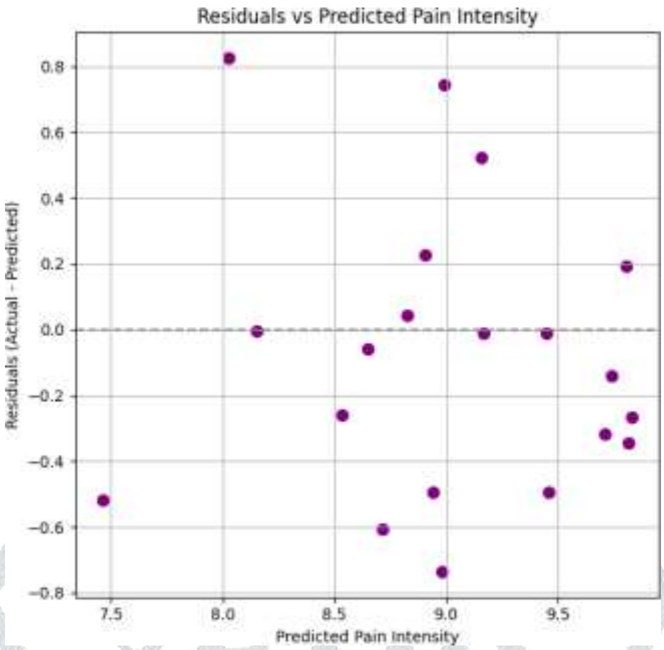

Fig. 1.2. Plot showing Residuals vs Predicted Pain Intensity of the model

In Figure 1.2, predicted pain intensity has been shown along x-axis, and the residuals (= Actual Pain Intensity – Predicted Pain Intensity) has been shown along the y-axis. Each point corresponds to one LBP patient in the test dataset. The figure does not show any distinct pattern, reflecting that the residuals are distributed almost symmetrically across the zero line. Most of the residual points are scattered within the range of +1 or -1, which indicates low prediction error.

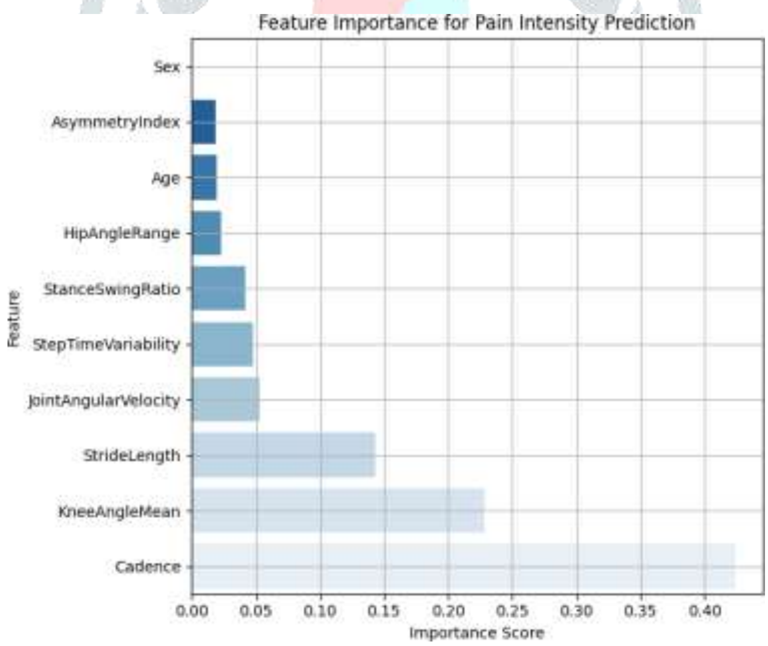

Fig. 1.3. Plot showing feature importance for prediction of pain intensity

In Figure 1.3, the gait kinematic parameters and the demographic parameters of the LBP patients have been plotted across the y-axis. The x-axis shows the feature importance scores that are calculated internally by XGBoost. The figure reflects that there are three most important gait parameters namely Cadence, Mean Knee Angle, and Stride Length for prediction of pain intensity.

Similar to the above plots, Figure 2.1 shows the plot to reflect Actual vs Predicted Pain Disability Score, Figure 2.2 shows the plot for Residuals vs Predicted Disability Score, and Figure 2.3 shows the plot for feature importance for prediction of pain-related disability.

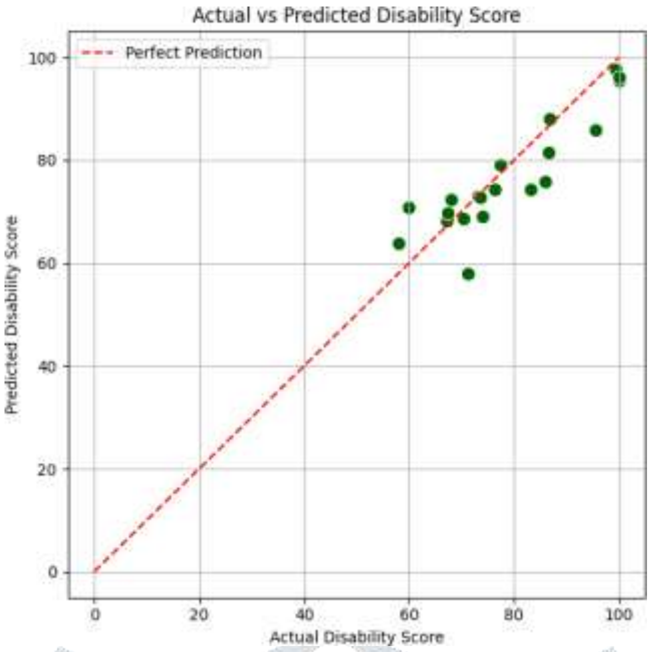

Fig. 2.1. Plot showing Actual vs Predicted Pain Disability Score of the model

In Figure 2.1, actual disability score has been shown along x-axis, and the predicted pain-related disability has been shown along the y-axis. Each green point corresponds to one LBP patient in the test dataset. Any point over the red diagonal line indicates that the prediction is exactly same as the actual pain related disability value. Figure 2.1 shows that there is tight clustering around the diagonal red line, implying that the prediction of pain related disability values are highly accurate. The presence of small deviations suggest model generalization and avoidance of large prediction errors. Low spreading of green points correspond to low MAE and low RMSE. No outliers are present. As the model shows high accuracy in prediction of pain related disability, it can be clinically useful.

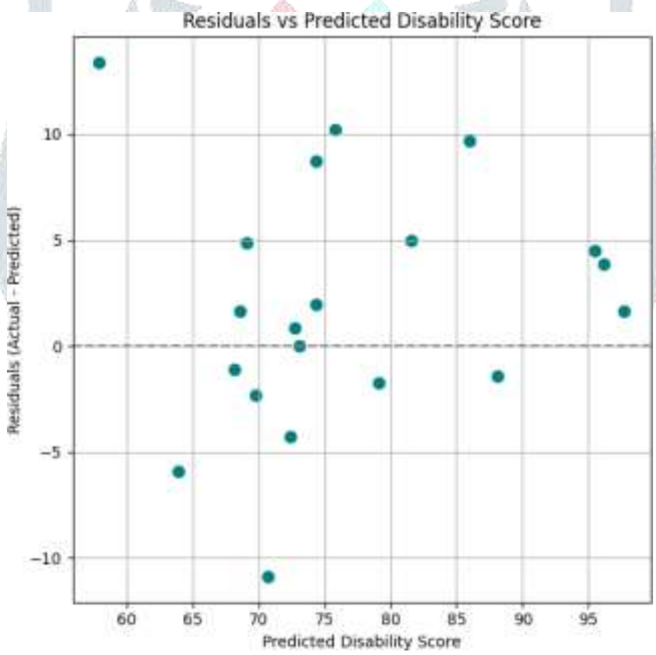

Fig. 2.2. Plot showing Residuals vs Predicted Pain Related Disability of the model

In Figure 2.2, predicted pain related disability has been shown along x-axis, and the residuals (= Actual Disability Score – Predicted Disability Score) has been shown along the y-axis. Each point corresponds to one LBP patient in the test dataset. The figure does not show any distinct pattern, reflecting that the residuals are distributed almost symmetrically across the zero line. Most of the residual points are scattered within the range of +10 or -10, which indicates low prediction error.

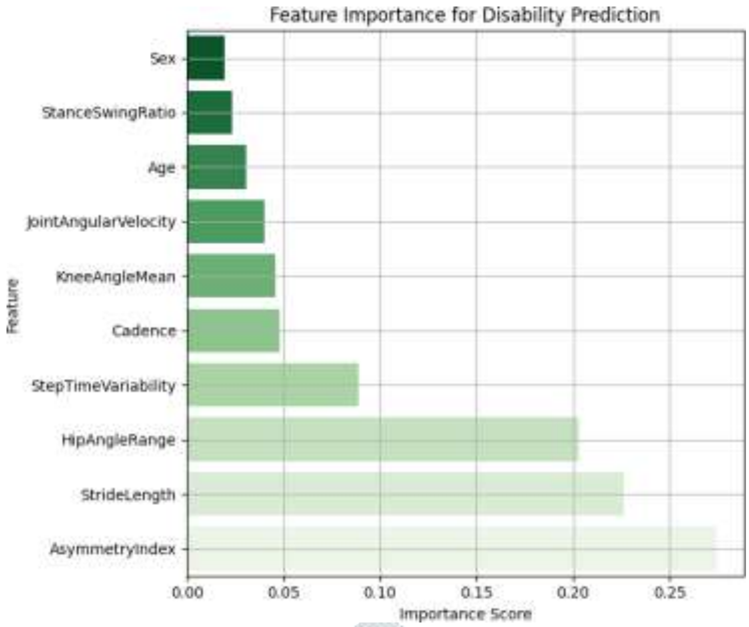

Fig. 2.3. Plot showing feature importance for prediction of pain related disability

In Figure 2.3, the gait kinematic parameters and the demographic parameters of the LBP patients have been plotted across the y-axis. The x-axis shows the feature importance scores that are calculated internally by XGBoost. The figure reflects that there are three most important gait parameters namely Asymmetry Index, Stride Length, and Hip Angle Range for prediction of pain related disability.

IV. DISCUSSION

The results of this study suggest that the proposed AI-based gait analysis model is very effective in estimating the pain intensity in patients with LBP. The fact that the value of the actual and predicted pain intensity is closely related as well as the low MAE and RMSE values indicate that the model is strong and can be generalized effectively with minimal errors in prediction. The small differences between estimated and measured values can be explained by the fact that the model is regularized, and it is capable of generalizing to the variability between inter-individual gait patterns. Taken together, these outcomes indicate high chances of clinical relevance of the suggested framework. Cadence, mean knee angle, and stride length were the most significant predictors of pain intensity of the gait kinematic parameters, which highlights the significance of temporal and lower-limb kinematic adjustments in the manifestation of pain.

In addition, the model showed a perfect match between the estimated and the measured disability scores in relation to the pain, which indicates that there was high predictive specificity of the model as far as functional impairment related to LBP is concerned. This excellent level of performance constitutes the clinical applicability of this framework of objective disability measurement. The range of the hip angle, the stride length, and the asymmetry index were found to be the most critical gait characteristics to predict the pain-related disability, which demonstrated the significance of bilateral coordination and hip–pelvic mechanics in the functionality limitation. Comprehensively, these results suggest the outcome of the proposed AI model as a valid, objective method of measuring both the intensity of pain and the disability caused by it, the results of which may be used in the specific planning of rehabilitation and monitoring of outcomes.

V. CONCLUSION AND FUTURE SCOPE

The study suggests that sagittal gait parameters can be used to objectively predict pain intensity and disability in LBP patients using machine learning. With disability scores calculated with low error on a 100-point scale and pain intensity predicted within ±1.0 units on a 10-point scale, the proposed model demonstrated great accuracy. The approach's clinical validity was strengthened by the discovery that important gait characteristics like cadence, joint angles, and stride length were strongly predictive. The suggested approach is particularly useful in clinical settings with limited resources or telemedicine since it provides a low-cost, non-invasive, and interpretable substitute for traditional pain assessment instruments.

Future research will concentrate on integrating 3D gait analysis for increased sensitivity, verifying the model with a large number of LBP patient data, and implementing the system on wearable or mobile platforms for real-time monitoring. The model's clinical relevance in diagnosis, treatment planning, and rehabilitation monitoring will be further improved by adding longitudinal tracking and expanding it to classify LBP subtypes.

Financial support and sponsorship: Nil.
Conflicts of interest: There are no conflicts of interest.

REFERENCES

[1] World Health Organization. Priority medicines for Europe and the world 2013 update: Chapter 6.6 Low back pain. Geneva: World Health Organization; 2013.

[2] Buchbinder R, van Tulder M, Öberg B, Costa LM, Woolf A, Schoene M, et al. Low back pain: A call for action. Lancet. 2018; 391(10137):2384–2388.

[3] Vos T, Abajobir AA, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016. Lancet. 2017; 390(10100):1211–1259.

[4] Dieleman JL, Baral R, Birger M, et al. US spending on personal health care and public health, 1996–2013. JAMA. 2016;

316(24):2627–2646.
[5] Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet. 2017; 389(10070):736–747.
[6] Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. Lancet. 2012; 379(9814):482–491.
[7] Koes BW, van Tulder M, Lin CW, et al. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. Eur Spine J. 2010; 19(12):2075–2094.
[8] Hoy D, March L, Brooks P, et al. The global burden of low back pain: Estimates from the Global Burden of Disease 2010 study. Ann Rheum Dis. 2014; 73(6):968–974.
[9] Ghamkhar L, Kahlaee AH, Khamseh M. Alterations in postural sway and gait parameters in patients with nonspecific chronic low back pain: A systematic review. Disabil Rehabil. 2015; 37(6):529–536.
[10] Hamacher D, Hamacher D, Schega L. Gait variability in chronic back pain sufferers with experimentally diminished visual feedback: A pilot study. Gait Posture. 2014; 39(1):235–238.
[11] Voloshina AS, Koldenhoven RM, Ferris DP. Biomechanics and energetics of walking and running with a load: A review. J Neuroeng Rehabil. 2018; 15:41.
[12] Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults. J Pain Symptom Manage. 2011; 41(6):1073–1093.
[13] Roland M, Morris R. A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low-back pain. Spine. 1983; 8(2):141–144.
[14] da Fonseca JL, Magini M, de Freitas TH. Laboratory gait analysis in patients with low back pain before and after a Pilates intervention. J Sport Rehabil. 2009; 18(2):269–282.
[15] Khodadadeh S, Eisenstein SM. Gait analysis of patients with low back pain before and after surgery. Spine. 1993; 18(11):1451–1455.
[16] Vickers J, Reed A, Decker R, et al. Effect of investigator observation on gait parameters in individuals with and without chronic low back pain. Gait Posture. 2017; 53:35–40.
[17] Chan H, Zheng H, Wang H, Sterritt R, Newell D. Smart mobile phone based gait assessment of patients with low back pain. In: Proc 9th Int Conf Natural Computation (ICNC). 2013:1062–1066.
[18] Mullerpatan R, Bharnuke J, Hiller C. Gait kinematics of Bharatanatyam dancers with and without low back pain. Crit Rev Phys Rehabil Med. 2019; 31(1).
[19] Baker R. Gait analysis methods in rehabilitation. J Neuroeng Rehabil. 2006; 3:4.
[20] Muro-De-La-Herran A, Garcia-Zapirain B, Mendez-Zorrilla A. Gait analysis methods: An overview of wearable and non-wearable systems. Sensors. 2014; 14(2):3362–3394.
[21] Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. IEEE Trans Pattern Anal Mach Intell. 2019; 43(1):172–186.
[22] Liao Y, Yang H, Zhang Y. Artificial intelligence in gait analysis: A systematic review. IEEE Access. 2022; 10:10107–10124.
[23] Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proc 22nd ACM SIGKDD Int Conf Knowl Discov Data Min. 2016:785–794.
[24] Friedman JH. Greedy function approximation: A gradient boosting machine. Ann Stat. 2001; 29(5):1189–1232.
[25] Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot. 2013; 7:21.
[26] Nielsen D. Tree boosting with XGBoost: Why does XGBoost win 'every' machine learning competition? Technical report. NTNU; 2016.