Support Vector Machine for Accurate Classification of Chronic Low Back Pain Severity

Marcelle Bolotari Fonseca Faciroli, Jair Moreira Dias Jr, Matheus Augusto Malta Ferreira, Nádia Rezende Barbosa Raposo, Eduardo Pestana de Aguiar

Submitted to: JMIR AI
on: October 24, 2024

Table of Contents

Support Vector Machine for Accurate Classification of Chronic Low Back Pain Severity

Marcelle Bolotari Fonseca Faciroli[1*]; Jair Moreira Dias Jr[2, 3*] MD, MSc; Matheus Augusto Malta Ferreira[4*] MSc; Nádia Rezende Barbosa Raposo[2*] PhD; Eduardo Pestana de Aguiar[5*] PhD

[1]Department of Eletrical Circuits School of Engineering Federal University of Juiz de Fora Juiz de Fora BR
[2]Center for Research and Innovation in Health Sciences (NUPICS) School of Pharmacy Federal University of Juiz de Fora Juiz de Fora BR
[3]Department of Orthopedics and Traumatology University Hospital Federal University of Juiz de Fora Juiz de Fora BR
[4]Electrical Engineering Graduate Program Federal University of Juiz de Fora Juiz de Fora BR
[5]Department of Mechanical Engineering Federal University of Juiz de Fora Juiz de Fora BR
[*]these authors contributed equally

Corresponding Author:
Marcelle Bolotari Fonseca Faciroli
Department of Eletrical Circuits
School of Engineering
Federal University of Juiz de Fora
Rua José Lourenço Kelmer, s/n - São Pedro
Juiz de Fora
BR

Abstract

This study applies Support Vector Machines (SVM) to enhance the diagnosis of chronic low back pain, achieving promising results and paving the way for more accurate and clinically impactful models.

(JMIR Preprints 24/10/2024:67951)
DOI: https://doi.org/10.2196/preprints.67951

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✔ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Research Letter

Marcelle B F Faciroli[1]; Jair M Dias Jr[2,3], MD, MSc; Matheus A M Ferreira[4], MSc; Nádia R B Raposo[2], PhD; Eduardo P de Aguiar[5], PhD
[1] Department of Eletrical Circuits, Federal University of Juiz de Fora, Juiz de Fora, Brazil
[2] Center for Research and Innovation in Health Sciences (NUPICS), School of Pharmacy, Federal University of Juiz de Fora, Juiz de Fora, Brazil
[3] Department of Orthopedics and Traumatology, University Hospital, Federal University of Juiz de Fora, Juiz de Fora, Brazil
[4] Electrical Engineering Graduate Program, Federal University of Juiz de Fora, Juiz de Fora, Brazil
[5] Department of Mechanical Engineering, Federal University of Juiz de Fora, Juiz de Fora, Brazil

Corresponding Author:
Marcelle B F Faciroli
Department of Eletrical Circuits
Federal University of Juiz de Fora
Rua José Lourenço Kelmer, s/n - São Pedro
Juiz de Fora, 36036-900
Brazil
Phone: +55 32 21023402
Email: marcelle.faciroli@estudante.ufjf.br

Support Vector Machine for Accurate Classification of Chronic Low Back Pain Severity

Abstract

This study applies Support Vector Machines (SVM) to enhance the diagnosis of chronic low back pain, achieving promising results and paving the way for more accurate and clinically impactful models.

Keywords: low back pain; support vector machine; SVM; machine learning; artificial intelligence

Introduction

Low back pain (LBP) is the leading cause of work-related disability worldwide, often classified as non-specific when no structural abnormalities are found. In contrast, radicular pain from neural compression is specific but affects a diverse group, making personalized treatment challenging. Current classification efforts aim to optimize treatment based on clinical data but need a widely accepted method [1].

Artificial intelligence (AI), particularly through machine learning (ML) techniques such as Support Vector Machines (SVM), offers promise by identifying subgroups among chronic LBP cases and categorizing diagnoses as either neural compression or non-specific pain. These AI-driven methods can significantly enhance clinical decision-making, providing more precise diagnoses and personalized treatment plans, ultimately improving patient outcomes [2,3].

Methods

Dataset

The dataset consisted of 168 real cases of LBP from patients who had a face-to-face consultation with a spine specialist and completed a survey with clinical and demographic questions, described in [4]. After the appointment, the doctor diagnosed either non-specific LBP or LBP with neural compression (LBPNC). The binary classification model applied the survey questions as the input variables (features) and the diagnosis as the output variable (labels).

This study was approved by the Research Ethics Committee of the University Hospital from the Federal University of Juiz de Fora, with registration number 69817223.7.0000.5133.

Preprocessing

Age, weight, and height columns were normalized to values between 0 and 1 [5]. Other features were preprocessed by creating dummy variables, where categorical variables were converted into a series of binary variables. This step was crucial as ML algorithms require numerical input. Multiple choice questions were split into new columns, one for each alternative, receiving a value of 1 or 0, based on the selected responses [4].

Fisher's Exact Test

The initial dataset had 38 features, which were expanded to 129 features after the dummy variables creation. Thus, it was necessary to perform an analysis to select attributes using Fisher's Exact Test (FET) [6]. This test was applied only to binary variables, whose frequencies were calculated and inserted into contingency tables with label frequencies.

By calculating the *P* value for each table using FET, we determined which features would reject the null hypothesis, with $P<.05$, as detailed in [4]. As a result, nine categorical features were retained, while the others were removed to prepare the dataset for ML modeling.

Support Vector Machine Model

SVM, a supervised learning algorithm, was chosen for its ability to handle high-dimensional data and small sample sizes, which are typical in clinical datasets. Its optimal margin maximization ensures reliable separation between classes, crucial for distinguishing non-specific LBP and LBPNC, making it an effective and interpretable model [7,8].

Two SVM models were implemented: one using the full dataset and the other using features selected by FET. Models' hyperparameters were tuned with grid search [9] and cross-validation [10] to mitigate the risk of overfitting and improve model generalization, and the best combinations were applied for model evaluation.

Results

The evaluated metrics are presented in Table 1, comparing the performance of models using all features and those with selected features. Since the models were trained and tested using 5-fold cross-validation, the results in Table 1 show each metric's mean and standard deviation across folds.

Table 1. Results of the SVM model.

	Accuracy	Precision	Recall	F1-Score	ROC-

					AUC
All features					
	0.702317±0.019656	0.666667±0.278887	0.082222±0.075556	0.124276±0.110577	0.574
Selected features					
	0.785383±0.049458	0.763333±0.198997	0.384444±0.109905	0.507143±0.131998	0.761

Discussion

This study proposed a classification approach between two types of LBP based on ML, data processing, and feature selection. While recognizing the relevance of all the metrics presented in Table 1, it is important to emphasize that accuracy is the most significant for this paper, as it shows the ratio of correct predictions to total predictions.

Therefore, the model trained with the FET chosen features outperformed the one with all initial features. It classifies cases of LBP more accurately than the literature results, emphasizing that it was fitted with fewer, but real samples, whereas the literature includes fictitious cases [2].

Future studies aim to increase the number of samples, refine survey questions, and make a comparison with other ML models, improving other clinically relevant metrics.

Conflicts of Interest

None declared.

References

1. Tagliaferri SD, Angelova M, Zhao X, Owen PJ, Miller C T, Wilkin T, Belavy DL. Artificial intelligence to improve back pain outcomes and lessons learnt from clinical classification approaches: three systematic reviews. NPJ Digit. Med. 2020;3:1-16. [doi: 10.1038/s41746-020-0303-x] [Medline: 32665978]
2. Oude Nijeweme-d'Hollosy W, van Velsen L, Poel M, Groothuis-Oudshoorn CGM, Soer R, Hermens H. Evaluation of three machine learning models for self-referral decision support on low back pain in primary care. Int J Med Inform. 2018;110:31-41. [doi: 10.1016/j.ijmedinf.2017.11.010] [Medline: 29331253]
3. Shim JG, Ryu KH, Cho EA, et al. Machine learning approaches to predict chronic lower back pain in people aged over 50 years. Medicina (Kaunas). 2021;57(11):1230. [doi: 10.3390/medicina57111230] [Medline: 34833448]
4. Faciroli, MBF. Support vector machine model for low back pain classification. GitHub. 2024. URL: https://github.com/marcellefaciroli/LowBackPain_classification.git [accessed 2024-10-08]
5. Singh D, Singh B. Investigating the impact of data normalization on classification performance. Applied Soft Computing. 2020;97(105524). [doi: 10.1016/j.asoc.2019.105524]
6. Fisher RA. On the interpretation of χ^2 from contingency tables, and the calculation of P.

Journal of the Royal Statistical Society. 1922;85(1):87-94. [doi: 10.2307/2340521]
7. Chang CC, Lin J. LIBSVM: A library for support vector machines. ACM Transactions Intelligent Systems and Technology (TIST). 2011;2(3):1-27. [doi: 10.1145/1961189.1961199]
8. Cervantes J, Garcia-Lamont F, Rodríguez-Mazahua L, Lopez A. A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing. 2020;408:189–215. [doi: 10.1016/j.neucom.2019.10.118]
9. Géron A. Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow. 2nd edition. Sebastopol, CA: O'Reilly Media. 2019. ISBN: 9781492032649
10. James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning: With Applications in Python. 2nd edition. New York, NY: Springer; 2023. ISBN:9783031387470 [Free Full Text]