Towards a Better Understanding of Spinal Differences Between Healthy Subjects and Subjects with Back Pain Using Explainable Artificial Intelligence (XAI)

Carlo Dindorf[1](✉), Jürgen Konradi[2], Claudia Wolf[2], Bertram Taetz[3], Gabriele Bleser[1], Eva Bartaguiz[1], Johanna Kniepert[2], Philipp Drees[2], Michael Fröhlich[1], and Ulrich Betz[2]

[1] Technische Universität Kaiserslautern, 67663 Kaiserslautern, Germany
carlo.dindorf@sowi.uni-kl.de
[2] Universitätsmedizin Johannes Gutenberg-Universität Mainz, 55131 Mainz, Germany
[3] German Research Center for Artificial Intelligence, 67663 Kaiserslautern, Germany

Abstract. Using Surface Topography data of stance measurements this study classifies 25 healthy subjects and 32 subjects with back pain using machine learning algorithms (Logistic Regression, Support Vector Machine, Random Forest). Different metric learning approaches (Neighborhood Components Analysis, Local Fisher Discriminant Analysis, Large Margin Nearest Neighbor) were applied to check if they lead to improved classification performance. Interpretations are performed using the Explainable Artificial Intelligence (XAI) tool SHapley Additive exPlanations (SHAP). The best results are obtained using Logistic Regression without prior performing a metric learning approach (MCC = 0.27, AUC = 0.71). Hence, a low correlation between predicted class and actual class is present. Results indicate that subjects with back pain exhibit a different posture than healthy subjects. The data driven approach could be useful to give clinicians and therapists an objective orientation and to individually adapt therapy measures. As a next step, the use of dynamic spinal data for classification should be evaluated.

Keywords: Spine · Back pain · Metric learning · Explainable artificial intelligence

1 Introduction

Surface Topography allows the measurement of the spine, both static and dynamic, without the usage of invasive, radiation-based approaches, or extensive preparation [1]. Back pain is of high social, clinical, and economic relevance. The extent to which healthy people can be distinguished from people with back pain based on spinal data is unclear [2]. Data driven approaches and the classification of pathologic characteristics proofed useful for giving an objective orientation and finding discriminative group specific differences. The usage of Explainable Artificial Intelligence (XAI) has especially shown

A. Baca et al. (Eds.): PACSS 2021, AISC 1426, pp. 97–100, 2022.
https://doi.org/10.1007/978-3-030-99333-7_16

to be useful in understanding individual pathologic differences and is therefore of high relevance in the context of personalized medicine [3]. Therefore, we want to check if classification of subjects with back pain and healthy subjects is possible and get insights into underlying biomechanical differences using XAI. Further, we want to check, if metric learning (mapping objects into an embedded space through learning a representation function) improves classification accuracy.

2 Methods

Static data of 25 healthy subjects (13 female, 12 male) and 32 subjects with back pain (18 female, 14 male) was recorded. For each subject, 12 recordings were used. 55 static parameters (for a detailed description see [4]) were used for modelling (Pelvic Obliquity [°], Pelvic Torsion (dimples) [°], Pelvic Inclination (dimples) [°], Pelvic Rotation [°], orientation of VP, T1-T12, L1-L4 in all planes [°]). Outliers were detected and removed using Isolation Forest algorithm (100 base estimators). Leave One Group Out Cross Validation was used for evaluation of three different classifiers (Logistic Regression with L1 regularization, Support Vector Machine with Radial Basis Function kernel, Random Forest with 100 trees). Standardization was performed based on the respective training set by removing the mean and scaling to unit variance. Due to imbalanced data, Synthetic Minority Oversampling Technique (SMOTE) was performed. Accuracy for the use of different metric learning approaches (Neighborhood Components Analysis, Local Fisher Discriminant Analysis, Large Margin Nearest Neighbor) was compared. SHapley Additive exPlanations (SHAP) was used as an XAI tool for model interpretation. Calculations were performed in Python (Python Software Foundation, Wilmington, DE, USA).

3 Results

The overall best results are obtained using Logistic Regression with L1 regularization without prior performing a metric learning approach (see Table 1; MCC = 0.27, AUC = 0.71). Looking at the most important features for the correctly classified samples shows that the features map characteristics in all body planes (sagittal plane: T3, T5, T6, VP; frontal plane: VP, T4, T6, T8; transverse plane: T3, T2). Figure 1 shows the top three features and their SHAP values. Looking at T3 in sagittal plane, values under approximately 20° indicate an effect towards the class of healthy, over20° for the class of subjects with back pain. For T3 in transverse plane, values under approximately 0° indicate the class of healthy, over 0° the class of subjects with back pain. The other way around, values under approximately 0° indicate an effect towards the class of subjects with back pain for VP in frontal plane. Values over approximately 0° indicate an effect for the class of healthy subjects. SHAP values can also be used for local interpretations.

Table 1. Classification results for the use of the three classifiers as well as the metric learning (ML) approaches. LR = Logistic Regression, SVM = Support Vector Machine, RF = Random Forest, NCA = Neighborhood Components Analysis; LFDA = Local Fisher Discriminant Analysis; LMNN = Large Margin Nearest Neighbor; AUC = Precision-Recall Area Under Curve, CM = Confusion Matrix

ML	LR				SVM				RF			
	MCC	AUC	CM		MCC	AUC	CM		MCC	AUC	CM	
–	0.27	0.71	194	106	0.05	0.60	128	172	0.01	0.55	128	172
			144	240			144	240			139	245
NCA	0.23	0.70	185	115	0.02	0.55	128	172	0.11	0.59	117	183
			146	238			157	227			111	273
LFDA	0.06	0.62	164	136	0.20	0.62	176	124	0.09	0.60	146	154
			188	196			148	236			153	231
LMNN	0.23	0.75	178	122	0.03	0.61	222	78	0.15	0.65	151	149
			140	244			275	109			135	249

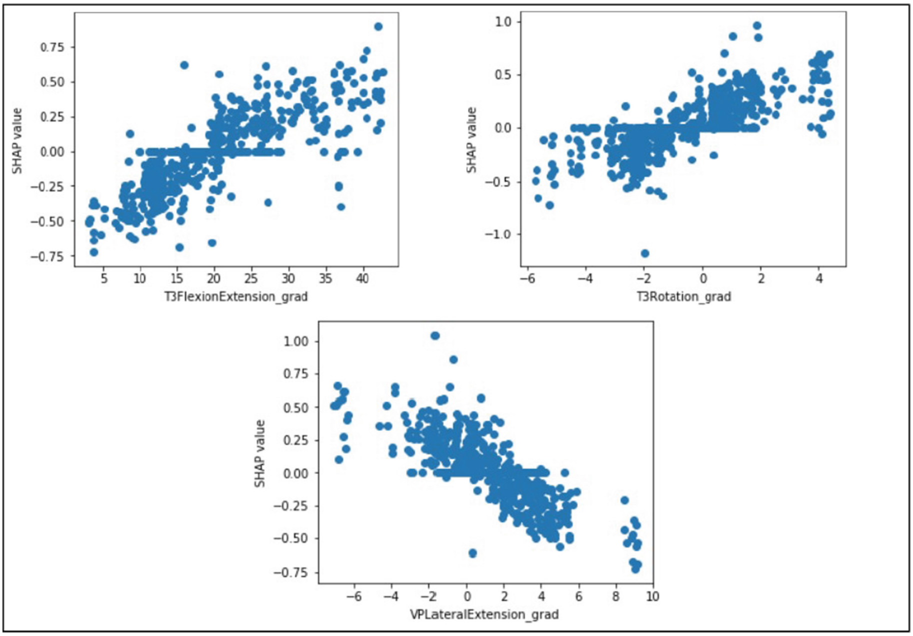

Fig. 1. SHAP values in dependence of the original feature values for the three most important features. A negative SHAP value indicates an effect towards the class of healthy subjects, a positive value an effect towards the class of subjects with back pain.

4 Discussion

Predicted class and actual class are only weakly correlated when using stance data. No total improvement of the performance is present using the current studies metric learning approaches. Results indicate that subjects with back pain exhibit a different posture than healthy subjects. However, the differences seem relatively small, which results in low discriminative power of the features. Interpretation of the results should therefore be done with caution.

Thoracic vertebrae and VP show the overall highest impact for the classification task. Healthy and subjects with back pain are therefore particularly different regarding the thoracic vertebrae and the VP according to the current study's findings.

The demonstrated data distribution for the class of healthy subjects and those with back pain (Fig. 1) shows that especially around the mentioned limit values, SHAP values are broadly spread and indicate an effect often for both classes. The more the feature values are distant from the limit value, the clearer the effect for a certain class.

Limitations arise through the relatively small sample of subjects. Therefore, differences could also be due to the sample and not due to actual pathologic differences between the groups. An expansion of the sample is therefore necessary for following studies.

5 Conclusion

The current data driven study indicates prevalent spinal differences between healthy subjects and subjects with back pain based on static data. As a next step, the use of dynamic spinal data should be evaluated. The data driven approach could be useful to give clinicians and therapists an objective orientation and to individually adapt therapy measures.

References

1. Dindorf, C., et al.: General method for automated feature extraction and selection and its application for gender classification and biomechanical knowledge discovery of sex differences in spinal posture during stance and gait. Comput. Methods Biomech. Biomed. Engin. **24**(3), 1–9 (2020). https://doi.org/10.1080/10255842.2020.1828375
2. Hildebrandt, J.: Funktionsund Leistungsdiagnostik. In: Hildebrandt, J., Pfingsten, M. (eds.) Rückenschmerz und Lendenwirbelsäule: Interdisziplinäres Praxisbuch entsprechend der Nationalen VersorgungsLeitlinie Kreuzschmerz, pp. 163–171. Urban & Fischer, München (2012). https://doi.org/10.1016/b978-3-437-23251-0.10001-x
3. Dindorf, C., Teufl, W., Taetz, B., Bleser, G., Fröhlich, M.: Interpretability of input representation for gait classification in patients after total hip arthroplasty. Sensors **20**(16), 1–14 (2020). https://doi.org/10.3390/s20164385
4. Dindorf, C., et al.: Classification and automated interpretation of spinal posture data using a pathology-independent classifier and explainable artificial intelligence (XAI). Sensors **21**(18), 1–16 (2021). https://doi.org/10.3390/s21186323