

Article

# Linear Support Vector Machines and Head Pitch Angular Velocity Are the Best Performers to Distinguish (Sub-)Acute Neck Pain Patients and Controls Assessed with the DidRen Laser Test

Renaud Hage<sup>1,2,\*</sup>, Fabien Buisseret<sup>1,3</sup>, Martin Houry<sup>4</sup>, Frédéric Dierick<sup>1,5,6</sup>

<sup>1</sup> CeREF Technique, Chaussée de Binche 159, 7000 Mons, Belgium ;

<sup>2</sup> Traitement Formation Thérapie Manuelle (TFTM), Private Physiotherapy/Manual Therapy Center, Avenue des Cerisiers 211A, 1200 Brussels, Belgium ;

<sup>3</sup> Service de Physique Nucléaire et Subnucléaire, UMONS, Research Institute for Complex Systems, Place du Parc 20, 7000 Mons, Belgium ;

<sup>4</sup> Haute-Ecole de Namur-Liège-Luxembourg (Henallux), Centre de Recherche FoRS, Rue Victor Libert 36H, 6900 Marche-en-Famenne, Belgium ;

<sup>5</sup> Centre National de Rééducation Fonctionnelle et de Réadaptation – Rehazenter, Laboratoire d'Analyse du Mouvement et de la Posture (LAMP), Rue André Vésale 1, 2674 Luxembourg, Luxembourg ;

<sup>6</sup> Faculté des Sciences de la Motricité, Université Catholique de Louvain, Place Pierre de Coubertin 1, 1348 Louvain-la-Neuve, Belgium.

\* Correspondence [renaudhage@gmail.com](mailto:renaudhage@gmail.com)

**Abstract:** Understanding neck pain is an important societal issue. Kinematic data from sensors may help to gain insight on the pathophysiological mechanisms associated with neck pain through a quantitative sensorimotor assessment of one patient. The objective of this study was to evaluate the potential usefulness of artificial intelligence with several Machine Learning (ML) algorithms in assessing neck sensorimotor performance. Angular velocity and acceleration measured by an inertial sensor placed on the forehead during the DidRen laser test in thirty-eight acute and subacute non-specific neck pain (ANSP) patients were compared to forty-two healthy control participants (HCP). Seven supervised ML algorithms were chosen for the predictions. The most informative kinematic features were computed using Sequential Feature Selection methods. The best performing algorithm is the Linear Support Vector Machine with an accuracy of 82% and Area Under Curve of 84%. The best discriminative kinematic feature between ANSP patients and HCP is the first quartile of head pitch angular velocity. This study has shown that supervised ML algorithms could be used to classify ANSP patients and identify discriminatory kinematic features potentially useful for the clinicians in the assessment and monitoring of the neck sensorimotor performance in ANSP patients.

**Keywords:** Artificial intelligence, Supervised Machine Learning; Kinematics; Head rotation test; Neck pain.

---

## 1. Introduction

Understanding neck pain is an important societal issue [1, 2]. The overall prevalence of neck pain in the general population ranges from 0.4% to 86.8% and is higher in women than in men [3]. It ranks fourth in terms of years lived with a disability [1, 2]. The majority of patients with neck pain are now classified as experiencing a "non-specific" neck disorder [4-6], meaning neck pain that occurs without trauma, signs or symptoms of major structural pathology, neurologic signs or specific pathology [4]. Acute or subacute non-specific neck pain (ANSP) means that the pain has been present for less than three months [4, 7]. The assessment of sensorimotor function, a generic term for tests that encompass all afferent and efferent information flows and central integration mechanisms that contribute to joint stability [8], has demonstrated its importance for a better understanding of the pathophysiological mechanisms associated with chronic neck pain [9]. Indeed, the assessment of sensorimotor function, especially through kinematics of the head rotations, seems promising for the identification of chronic neck pain [10]. Nevertheless, identification based on sensorimotor evaluation requires the ability to know what would characterize neck pain in terms of the kinematic features of movement. Sensorimotor assessment of neck motion based not only on position degrees of freedom but also on velocity and acceleration features (e.g. peak and average velocity) appears promising because it has high sensitivity and specificity [10, 11].

Identifying kinematic features from time series and comparing them between groups, e.g., to evaluate treatments or classify neck pain motion across ageing, is a widely used method [11-15]. Here we focus on a peculiar test called DidRen laser test, designed to assess sensorimotor control of the neck and about which the interested reader will find detailed information in [14-16]. The DidRen laser test consists of a standardized task in which yaw rotations of the head are performed from "target to target" in the same sequence. These are fast, accurate, and small-amplitude rotations ( $\pm 30^\circ$ ) of the head in response to real visual targets to be hit by a laser beam placed on the subject's head [17]. However, such a methodology removes a significant amount of information from the raw time series. This could lead to false-negative results in statistical inference. Resorting to artificial intelligence (AI) techniques may lead to another type of analysis, i.e., "the machine" should find the relevant specifications of time series. The present study is devoted to the latter case.

AI is defined as a field of science and engineering concerned with the computational understanding of what is commonly referred to as intelligent behavior and the creation of artefacts that exhibit such behavior [18]. Machine learning (ML) is defined as a subfield of AI as follows: "Machine learning is a branch of artificial intelligence that systematically applies algorithms to synthesize the underlying relationships among data and information" [19]. ML provides an experiential "learning" that can be related to human intelligence as ML can improve its analyses by using computer algorithms. There are two main forms of ML: supervised and unsupervised [20]. In supervised ML (SML), the algorithms are provided with training data that are analyzed for the features that are important for classification and labelled. The model is then "trained" on this data before being tested on unlabeled data. In our case, the data will be measured in head rotations. In SML, data must be first labelled by a clinician (painful or not for example) so that the model can learn to interpret them through pattern recognition. Then, the model is tested with unlabeled data to obtain an interpretation result [21, 22]. Several algorithms can be trained for pattern recognition such as logistic regression, support vector machine, decision tree, random forest, naïve Bayes or K-nearest neighbor [20]. Patterns may be representative of various features, among which pathology and pain, see e.g. [23].

The first aim of this work was to evaluate the discriminative ability of AI and SML methods in sensorimotor assessment of yaw angular displacement of the head in patients with ANSP patients compared with healthy control participants (HCP) with data from a previous study [15] obtained during the DidRen laser test [14, 16, 17]. A second aim of this work was to illustrate the potential of SML for clinicians in musculoskeletal physiotherapy [24]. In ecological situations, neck kinematics should be quickly assessed by a therapist using thresholds designed to identify relevant impairments in the history of patients with neck pain. We test whether SML can provide such kinematic values and therefore has predictive value for ANSP.

## 2. Materials and Methods

### 2.1. Patients and participants

This study included 80 subjects (38 ANSP patients and 42 HCP) from a previous study [15]. Data were collected from February to December 2019. ANSP patients diagnosed by general practitioners were recruited from a consecutive sample in a private manual therapy center in Brussels, Belgium. Inclusion criteria for ANSP patients were acute-subacute (< 3 months) non-specific neck pain with a Neck Disability Index (NDI)  $\geq$  8% [25] and a Numeric Pain Rating Scale (NPRS)  $>$  3 [26-30]. HCP were recruited by one of the authors (RH) from a sample of convenience from colleagues at the university hospital and from acquaintances. They were included if they reported no neck symptoms: NDI  $<$  8% [25], NPRS = 0 [26], and no pain on active head rotation and/or manual spinal assessment [31]. Characteristics of the ANSP patients and HCP are listed in Table 1. All subjects signed an informed consent form. The study was approved by the Academic Bioethics Committee (<https://www.a-e-c.eu>, Brussels, B200-2018-103) and conducted in accordance with the Declaration of Helsinki. The authors confirm that all ongoing and related trials for this drug/intervention are registered ([ClinicalTrials.gov: 04407637](https://clinicaltrials.gov/04407637)).

Table 1 Characteristics of the acute and subacute non-specific neck pain (ANSP) patients and healthy control participants (HCP). *P*-values resulted from t-test for age and BMI, Mann-Whitney U-test for NDI and NPRS, and Chi-2 for gender.

	ANSP patients (n=38)	HCP (n=42)	<i>P</i> -values
Age (years), mean $\pm$ SD	46.2 $\pm$ 16.3	24.3 $\pm$ 6.8	< 0.001
Gender n (men/women), (%)	21 (55%)/17 (45%)	27 (64%)/15 (36%)	0.55
BMI (kg m <sup>-2</sup> ), mean $\pm$ SD	23.5 $\pm$ 3.2	21.5 $\pm$ 4.2	0.014
NDI (100), median [Q1-Q3]	22 [16-31.5]	0 [0-0]	< 0.001
NPRS, median [Q1-Q3]	6 [4-7]	0 [0-0]	< 0.001

BMI: body mass index, NDI: neck disability index, NPRS: numeric pain rating scale

### 2.2. Protocol

The protocol was described in a previous study [15]. It essentially involved assessment of fast neck yaw rotations with the DidRen laser test [14, 16] for ANSP patients and HCP, completed by manual examination of the painful spinal region for segmental tenderness. For ANSP patients, the manual

examination served to confirm familiar pain and guide the treatment. For HCP, thanks to its high sensitivity (92%), the manual examination was used to exclude HCP if they had pain at one or more levels of the cervical spine and confirm that they are not healthy in the neck [31]. The DidRen laser test was used to standardize the rotational yaw movements of the participant's head. Briefly, participants wore a helmet to which a laser was attached. They pointed the laser as fast as possible at three targets equipped with photosensitive sensors (Fig 1A,B). The angular separation of targets is 30°, and the sequence was fixed: center-left-center-right-center. Participants were asked to perform the sequence as fast as possible.

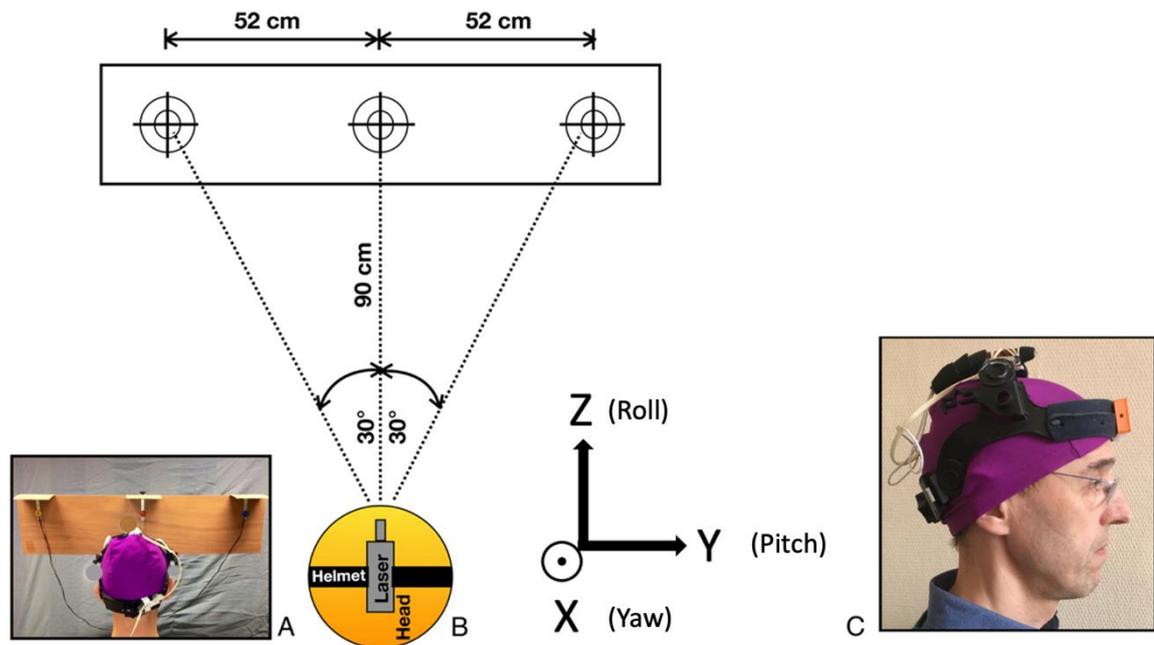


Fig 1. Description of the DidRen laser test. (A) Rear view of head position in front of the targets. (B) Schematic top view of the experimental setup with the three photosensitive sensors. The reference frame of the sensor is displayed when the head is in rest position. Coordinates system used in the study is also shown with the yaw (X-axis), pitch (Y-axis), and roll (Z-axis) rotations of the head during the test. (C) Helmet worn by an HCP (here RH) with laser on the top of the head and DYSKIMOT inertial sensor on the forehead.

During the DidRen laser test, head angular displacement kinematics was recorded in 3D (yaw, pitch, and roll) using the DYSKIMOT inertial sensor [32]. The detailed description of the sensor can be found in the study by Hage et al. [32]. The sensor consists of a 3-axis accelerometer, a gyroscope and magnetometer, and a temperature sensor. These internal components respectively measure acceleration (in g,  $\pm 16$  g), angular velocity (in  $^{\circ}/s$ ,  $\pm 2000$   $^{\circ}/s$ ), and magnetic field (in gauss,  $\pm 16$  gauss). The sensor recorded the motion at a sampling frequency of 100 Hz. The DYSKIMOT sensor was placed in front of the helmet (Figure 1C), with the yaw-axis (or X) in the vertical direction. The pitch-axis (or Y) was aligned with subject's medio-lateral axis at the start of the test and the roll-axis (or Z) was aligned with the antero-posterior axis. The head rotation demanded in the DidRen laser test is oriented along the yaw-axis. Note that the subjects were not instructed to realize pitch or roll rotations of the head during the test.

## 2.3 Data analysis

### 2.3.1 Dataset and pre-processing

The dataset consists of 7 time series for each participant: time, angular velocity (three components labelled GyrX, GyrY, GyrZ), and acceleration (AccX, AccY, AccZ). Then, a pre-processing procedure was applied to convert each time series into a summary format for all participants. Each time series is summarized with 7 statistical descriptors: 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> quartiles, mean, minimum, maximum, and standard deviation. The result is a dataset with 186 inputs and 42 features (6 time series x 7 descriptors). Each set of statistical descriptors is labeled as ANSP (value 1) or HCP (value 0).

### 2.3.2 ML algorithms and determination of the best performer

It is generally difficult to determine a priori which ML algorithm performs best on a given dataset [33]. Therefore, several algorithms were tested to determine the most appropriate for classifying ANSP patients and HCP: K-Nearest Neighbor (KNN), Linear Support Vector Machine (Linear SVM), Non-linear Support Vector Machine Radial Basis Function (SVM RBF), Decision Tree (DT), Random Forest (RF), Adaptive Boosting (AdaBoost), and Gaussian Naive Bayes (GaussianNB).

The comparison between selected algorithms was based on metrics such as accuracy and the Area Under Curve (AUC) score, computed from the Receiver Operating Characteristic (ROC) curve. These metrics are only meaningful if the predictions are based on data that the ML algorithms have never learned. Therefore, the dataset was randomly split into two parts. The first part is the “training set”, which consists of 80% of the data set used to train the ML algorithms. The second part (remaining 20%) is the “test set” used to make the predictions with the trained ML algorithms. The validation of the ML algorithms is done by n-fold cross-validation [34]. To minimize the biases associated with the training dataset, 100 different cross-validations were performed on mixed data for each selected ML algorithm. The hyperparameters of the ML algorithms were optimized using the Grid Search method [35] that finds the best combination of fixed hyperparameters based on n-fold cross-validation.

For KNN, the optimized parameters were the following: the number of neighbors (n\_neighbors: 3, 5, 8, 10), the weighting function (weights: uniform, distance) and the algorithms used to compute the nearest neighbors (algorithms: Brute-Force (BF KNN or BF KNN), kd\_tree, auto, ball\_tree). For Linear SVM, different values for the regularization parameter or C-parameter (0.1, 1, 10, 100, 1000) were used in the evaluation to test the dependence of the approach on the C-parameter. For SVM RBF, the C-parameter (0.001, 0.01, 0.1, 1, 10, 100) and the kernel coefficient Gamma (0.001, 0.01, 0.1, 1, 10, 100) parameter were optimized. For DT, the optimized parameters were the maximum depth of the tree (max\_depth: 1, 5, 10, 100), the function to measure the quality of the splits (criterion: gini, entropy) and the strategy to select the split nodes (splitter: best, random). For RF, the optimized parameters were the maximum depth of the tree (max\_depth: 1, 5, 10, 100), the number of trees in the forest (n\_estimators: 1, 5, 10, 100), the number of features considered in the search for the best split (max\_features: 1, 5, 10, 100). For Adaboost, the optimized parameters were the maximum number of estimators at which boosting stops (n\_estimators: 1, 5, 10, 50, 100, 500) and the weight applied to each classifier at each boosting iteration (learning\_rate: 0.000001, 0.001, 0.1, 1, 5, 10, 100). For GaussianNB, the optimized parameter was the ratio of the largest variance of all features added to the variances for computational stability (var\_smoothing: 0.0000001, 0.01, 1, 10, 100).

All the computations related to the determination of the best performer were made in Python 3.8 and SciKit-Learn 1.0.2 software.

### 2.3.2 Determination of most informative kinematic features and logistic regressions

The most informative kinematic features, i.e., the features that trigger the most predictions, were computed by using the Sequential Feature Selector (SFS) forward and backward [36]. The backward SFS removes the poorest features one by one, while the forward SFS identifies the best combination of features. In both cases, the result is a list of kinematic features that performed best according to the AUC score. Each SFS was run 700 times (7 ML algorithms x 100 random data repartitions). Once the most informative kinematic feature was identified, a logistic regression was performed by using it, and the accuracy of this logistic regression was computed. Another logistic regression on total DidRen laser test duration was also performed to compare the present results to the unique outcome of original DidRen laser test [17].

All the computations related to the determination of the most informative features and ML algorithms were made in Python 3.8 and SciKit Learn 1.0.2 software.

## 3. Results

### 3.1 Optimal hyperparameters and performance metrics of ML algorithms

Optimal hyperparameters are presented in Table 2. Performance metrics of the selected ML algorithms are given in Table 3. The least performing ML algorithm is the KNN, and the best performing one is the linear SVM with an accuracy of 82 % and AUC of 84%. We show in Figure 2 the ROC curve of the Linear SVM, which is the best ML algorithm we found to classify ANSP patients and HCP.

Table 2: Optimal hyperparameter values: Number of neighbors ( $n\_neighbors$ ), Regularization parameter (C-parameter), Kernel coefficient ( $\gamma$ ), maximum depth of the tree ( $max\_depth$ ), number of trees in the forest ( $n\_estimators$ ), and number of features to consider when looking for the best split ( $max\_features$ ).

ML algorithm	Hyperparameters
BF KNN	$n\_neighbors=5, weights="distance"$
Linear SVM	$kernel="linear", C=10$
SVM RBF	$\gamma=0.001, C=100$
DT	$max\_depth=1, criterion="entropy", splitter="best"$
RF	$max\_depth=10, n\_estimators=100, max\_features=10$

BF KNN: Brute-Force K-Nearest Neighbors, SVM: Support Vector Machine, RBF: radial basis function, DT: Decision Tree, RF: Random Forest

Table 3: Performance metrics of the selected ML algorithms.

ML algorithm	Accuracy	AUC score
BF KNN	0.66 ± 0.03	0.51 ± 0.07
Linear SVM	0.82 ± 0.03	0.84 ± 0.04
SVM RBF	0.65 ± 0.05	0.57 ± 0.09
DT	0.74 ± 0.03	0.70 ± 0.04
RF	0.76 ± 0.03	0.76 ± 0.04
AdaBoost	0.75 ± 0.04	0.76 ± 0.05
GaussianNB	0.77 ± 0.03	0.82 ± 0.03

BF KNN: Brute-Force K-Nearest Neighbors, SVM: Support Vector Machine, RBF: radial basis function, DT: Decision Tree, RF: Random Forest, AdaBoost: Adaptive Boosting, GaussianNB: Gaussian Naive Bayes, AUC: area under curve

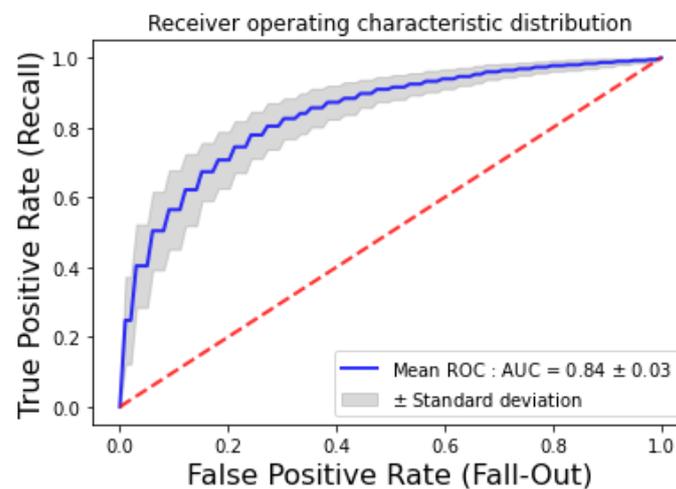


Figure 2: Receiver Operating Characteristic (ROC) curve of Linear SVM (in blue). The dotted red line represents the worst possible scenario, a random classifier.

### 3.2 Most discriminative features and logistic regressions

The most discriminative feature, regardless of the ML algorithm and SFS, was the 1<sup>st</sup> quartile of head pitch angular velocity (or GyrY), which ranked first 813 times on 1400. The second most discriminative feature was the median of head pitch angular velocity (ranked first 444 times on 1400). Thus, the pitch angular velocity appears to be the best discriminating feature to differentiate ANSP patients and HCP assessed with the DidRen laser test.

A logistic regression based on the median of head pitch angular velocity led to an accuracy of 77%. A logistic regression based on total duration of the DidRen test led to an accuracy of 63%.

## 4. Discussion

Our findings showed the effectiveness of the kernel linear SVM classifier in distinguishing ANSP patients from HCP. The accuracy of the linear SVM was 82% and the AUC score was 84%. The interpretation of the AUC score should be evaluated in terms of the importance given to its accuracy. We can assume that the medical community in the field of oncology prefers an AUC score close to 100%. Considering, on the one hand, the musculoskeletal field and, on the other hand, in relation to the non-specific pathology, the comparison between ANSP patients and HCP, which shows a great variability of the results [37], an AUC score higher than 80% can be considered satisfactory.

Seven time series (time and kinematic data) related to yaw, pitch, and roll angular displacement and velocity of the head, which can be easily acquired with a single inertial sensor, were used to train the selected ML algorithms. However, regardless of the ML algorithm and SFS, not all axes of head motion have good discriminative information, as the two best discriminating kinematic features were related to head pitch. The accuracy was best with the linear SVM and lowest with all other selected ML algorithms, such as the non-linear SVM (RBF). The same finding regarding the superiority of linear SVM over RBF has already been observed in a study with limited sample size (17 young and 17 old subjects) aimed at detecting age-related changes in running kinematics [38]. For use in future clinical trials with kinematic variables with limited sample size, linear SVM may thus be a suitable option.

Like other studies using ML algorithms to detect kinematic changes in healthy or pathological subjects [38-41], our study is based on a rather small dataset in terms of typical AI calculations, but the results are consistent with the conclusions of [42]. While conducting observational sensorimotor assessment studies with large datasets holds promise for improving the understanding and management of various pathologies, here, the pathophysiological mechanisms associated with neck pain, the use of small datasets may also allow for a reduction in selection bias [42]. In addition, it is worth noting that an SVM has already been used in the musculoskeletal field to compare temporomandibular patients with control subjects [43]. With a smaller sample (10 patients and 10 control subjects), they achieved an average predictive accuracy of 60% ( $P = 0.10$ ) [43]. The linear SVM algorithm is affordable with today's standard devices: a tablet computer could efficiently post-process the data from any wearable inertial sensor. Note also that a logistic regression based on head pitch angular velocity could be easily implemented on any smartphone, but with a lower accuracy of 77%.

The main discriminatory information used by the linear SVM algorithm to distinguish ANSP patients from HCP are the first quartile and the median of head pitch angular velocity. The best two kinematic discriminating features differed from those obtained by inferential statistical analysis, suggesting that ML approaches are complementary and clinically useful to detect kinematic changes in patients with ANSP.

HCP have larger medians and quartiles for head pitch angular velocity (computed from GyrY time series) than ANSP patients, making the Y-axis a highly discriminatory direction that should be prioritized for future clinical trials with the DidRen laser test. Our results may seem counterintuitive at first, because the DidRen laser test consists of a sensorimotor assessment organized around the Z-axis, i.e., during the execution of yaw rotations of the head. Thus, it would stand to reason that the GyrX time series should contain most of the discriminative information. Nevertheless, it is interesting to note that the sensorimotor disturbances in ANSP patients may be highlighted by the stronger

secondary coupled motion during yaw rotations. There seems to be a reason for this, because biomechanically, coupled bending rotations in the cervical spine lead to a compensatory roll rotation, which compensates for the yaw rotation of the head, and the associated coupled movements observed during pitch head movements [44]. Indeed, in HCP, we can observe that yaw head rotation ( $55.5 \pm 10.8^\circ$ ) is coupled with a larger pitch motion ( $16.3 \pm 11.4^\circ$ ) than roll motion ( $4.6 \pm 6.2^\circ$ ) [44]. If we apply these considerations to patient assessment, this information may be of clinical interest because 3D motion analysis may be a useful tool for assessing postural changes in the cervical spine during sitting, but also because altered kinematics are associated with decreased performance, e.g., neck velocity and neck motion fluidity in functional movement tasks, in people with neck pain [45].

The present discussion suggests that the ML algorithms can provide relevant functional variables and thus optimize the prediction of ANSP status during the DidRen laser test. To further illustrate this point, we mention that total test duration was the only parameter measured in the original version of the DidRen laser test [17]. Logistic regression performed with duration yields lower accuracy than that obtained with the median of pitch head rotation alone, the latter parameter being favored by linear SVM.

In experimental studies, control and experimental groups are usually formed in such a way that no significant difference is observed in parameters such as age, ethnicity, gender, and degeneration/maturation stage, except for the variable of interest. In our case, this means that ANSP patients and HCP groups should differ only in terms of NDI and NPRS. Age is also significantly different in our groups, but we do not believe this is problematic for our purpose. Indeed, ML algorithms are designed to distinguish between HCP and ANSP patients. To find out the characteristics of ANSP patients, it is logical to compare them with the “healthiest” subjects, i.e., our HCP group. On the other hand, a control group with too young subjects would also have led to bias, since we have shown in a previous study that the kinematic behaviors recorded with the DidRen laser test have a U-shaped or inverted U-shaped age profile, making the differences between young and old particularly clear [16]. Because the prevalence of degenerative joint changes increases with age [4, 46], possibly leading to movement limitations (yaw rotation steadily decreases between the ages of 30 and 60) [47], we selected HCP using a very sensitive manual examination [31]. After this examination, positive control subjects (with potential neck disorders) were excluded, and because their average age was higher (see [15]: the mean age of the excluded control subjects was 43.3 years), the average age of our HCP group decreased compared with the ANSP group. It is worth noting that a significant age difference between control and disease subjects was already found in a study aimed at developing and determining the predictive performance of ML models to distinguish between different subtypes of low-back pain and healthy control subjects [48]. For this purpose, as we did, they did not include age as a predictor when constructing the model [48].

We previously found several differences between ANSP patients and HCP on several features using inferential statistical analysis [15]: overshoot, stabilization time, peak speed, time to peak speed, average speed, peak acceleration, time to peak acceleration, peak deceleration, time to peak deceleration, time between peaks of acceleration and deceleration, time from peak acceleration to end of rotation, and angle at maximum speed. Such variables may be seen as relevant classifiers. For example, Sarig Bahat et al. [10] showed a sensitivity for right and left yaw rotations of the head of 91% and 94%, respectively, when calculating peak speed. Roijejon et al. [11] found a sensitivity of

74.6% for peak head speed rotation. For other variables, such as the number of peak speeds, Bahat et al. determined a sensitivity for right and left head yaw rotations of 73% and 85%, respectively [10]. Roijezon et al. determined a sensitivity for the normalized peak speed amplitude of 59.3% [11]. Here, our aim was to extend the inferential statistical analysis of [15] with selected ML classifier models. Given the responses to the patient data calculated using a classical statistical approach, it is obvious that the neck rotation task was performed with a rather important variability in the participants. For example, the peak speed was 88.06 (79.70-105.77) °/s in ANSP patients and 109.73 (87.84-128.58) °/s in HCP (written under the form median [Q1-Q3]). This may explain the different values of sensitivity between, our study and [11]. The difference between these values can be explained, of course, by the different evaluation methodology used in the these studies, but mainly, and this characterizes our approach, by the conversion of functional variables into scalar variables, which removes a significant part of information contained in the original variables [49]. In statistical inference, this can theoretically lead to results that do not correspond to reality [49]. Indeed, this could lead to false negative results or to a lack of precision in the prediction [49].

In summary, we have shown that AI can help identify patients suffering from neck pain using the DidRen laser test augmented by an inertial sensor. AI is therefore clinically useful information that can improve patient management. It can either predict the patient's condition based on the time series of angular velocity and acceleration or suggest the kinematic features most likely to distinguish neck pain from healthy individuals. In a sense, then, AI can suggest improvements to existing clinical tests.

**Author Contributions:** Conceptualization, R.H., F.D.; methodology, R.H., M.H.; software, R.H., M.H.; validation, R.H., F.B., F.D. and M.H.; data curation, R.H., M.H.; writing—original draft preparation, R.H., M.H.; writing—review and editing, R.H., F.B., F.D., and M.H.. All authors have read and agreed to the published version of the manuscript.

**Funding:** CeREF acknowledges financial support from the First Haute-Ecole program, project n° 1610401, DYSKIMOT, in partnership with OMT-Skills (<http://omtskills.be/>), and from the European Regional Development Fund (Interreg FWVI NOMADe).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by Academic Bioethics Committee (<https://www.a-e-c.eu/>) under the number B200-2018-103. The authors confirm that all ongoing and related trials for this drug/intervention are registered (ClinicalTrials.gov: 04407637).

**Informed Consent Statement:** Informed consent was obtained from all participants involved in the study.

**Data Availability Statement:** Data are available at <https://osf.io/rsuh2/>

## References

1. Vos T, Barber RM, Bell B, Bertozzi-Villa A, Biryukov S, Bolliger I, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*. 2015;386(9995):743-800. doi: 10.1016/s0140-6736(15)60692-4.
2. Hoy DG, Protani M, De R, Buchbinder R. The epidemiology of neck pain. *Best Pract Res Clin Rheumatol*. 2010;24(6):783-92. Epub 2011/06/15. doi: 10.1016/j.berh.2011.01.019. PubMed PMID: 21665126.

3. Kazeminasab S, Nejadghaderi SA, Amiri P, Pourfathi H, Araj-Khodaei M, Sullman MJM, et al. Neck pain: global epidemiology, trends and risk factors. *BMC Musculoskelet Disord.* 2022;23(1):26. Epub 2022/01/05. doi: 10.1186/s12891-021-04957-4. PubMed PMID: 34980079.
4. Blanpied PR, Gross AR, Elliott JM, Devaney LL, Clewley D, Walton DM, et al. Neck Pain: Revision 2017. *J Orthop Sports Phys Ther.* 2017;47(7):A1-a83. Epub 2017/07/02. doi: 10.2519/jospt.2017.0302. PubMed PMID: 28666405.
5. Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner RS, Whitman JM, et al. Neck pain: Clinical practice guidelines linked to the International Classification of Functioning, Disability, and Health from the Orthopedic Section of the American Physical Therapy Association. *J Orthop Sports Phys Ther.* 2008;38(9):A1-a34. Epub 2008/09/02. doi: 10.2519/jospt.2008.0303. PubMed PMID: 18758050.
6. Coulter ID, Crawford C, Vernon H, Hurwitz EL, Khorsan R, Booth MS, et al. Manipulation and Mobilization for Treating Chronic Nonspecific Neck Pain: A Systematic Review and Meta-Analysis for an Appropriateness Panel. *Pain physician.* 2019;22(2):E55-e70. Epub 2019/03/30. PubMed PMID: 30921975.
7. Pool JJ, Ostelo RW, Knol D, Bouter LM, de Vet HC. Are psychological factors prognostic indicators of outcome in patients with sub-acute neck pain? *Man Ther.* 2010;15(1):111-6. Epub 2009/09/01. doi: 10.1016/j.math.2009.08.001. PubMed PMID: 19717327.
8. de Zoete RMJ, Osmotherly PG, Rivett DA, Farrell SF, Snodgrass SJ. Sensorimotor Control in Individuals With Idiopathic Neck Pain and Healthy Individuals: A Systematic Review and Meta-Analysis. *Arch Phys Med Rehabil.* 2017;98(6):1257-71. Epub 2016/10/25. doi: 10.1016/j.apmr.2016.09.121. PubMed PMID: 27771360.
9. Sjolander P, Michaelson P, Jaric S, Djupsjobacka M. Sensorimotor disturbances in chronic neck pain--range of motion, peak velocity, smoothness of movement, and repositioning acuity. *Man Ther.* 2008;13(2):122-31. Epub 2007/01/02. doi: 10.1016/j.math.2006.10.002. PubMed PMID: 17197230.
10. Sarig Bahat H, Chen X, Reznik D, Kodesh E, Treleaven J. Interactive cervical motion kinematics: sensitivity, specificity and clinically significant values for identifying kinematic impairments in patients with chronic neck pain. *Man Ther.* 2015;20(2):295-302. Epub 2014/12/03. doi: 10.1016/j.math.2014.10.002. PubMed PMID: 25456272.
11. Roijezon U, Djupsjobacka M, Bjorklund M, Hager-Ross C, Grip H, Liebermann DG. Kinematics of fast cervical rotations in persons with chronic neck pain: a cross-sectional and reliability study. *BMC Musculoskelet Disord.* 2010;11:222. Epub 2010/09/30. doi: 10.1186/1471-2474-11-222. PubMed PMID: 20875135; PubMed Central PMCID: PMC3224737.
12. Sarig Bahat H, Weiss PL, Laufer Y. The effect of neck pain on cervical kinematics, as assessed in a virtual environment. *Arch Phys Med Rehabil.* 2010;91(12):1884-90. Epub 2010/11/30. doi: 10.1016/j.apmr.2010.09.007. PubMed PMID: 21112430.
13. Bahat HS, Igbariya M, Quek J, Treleaven J. Cervical Kinematics of Fast Neck Motion across Age. *Journal of Novel Physiotherapies.* 2016;6(5). doi: 10.4172/2165-7025.1000306.
14. Hage R, Dierick F, Roussel N, Pitance L, Detrembleur C. Age-related kinematic performance should be considered during fast head-neck rotation target task in individuals aged from 8 to 85 years old. *PeerJ.* 2019;7. doi: 10.7717/peerj.7095.
15. Hage R, Detrembleur C, Dierick F, Brismée JM, Roussel N, Pitance L. Sensorimotor performance in acute-subacute non-specific neck pain: a non-randomized prospective clinical trial with intervention. *BMC Musculoskelet Disord.* 2021;22(1):1017. Epub 2021/12/06. doi: 10.1186/s12891-021-04876-4. PubMed PMID: 34863120.

16. Hage R, Buisseret F, Pitance L, Brismee JM, Detrembleur C, Dierick F. Head-neck rotational movements using DidRen laser test indicate children and seniors' lower performance. *PLoS One*. 2019;14(7):e0219515. Epub 2019/07/26. doi: 10.1371/journal.pone.0219515. PubMed PMID: 31344044.
17. Hage R, Ancenay E. Identification of a relationship between cervical spine function and rotational movement control. *Ann Phys Rehabil Med*. 2009;52(9):653-67. Epub 2009/09/29. doi: 10.1016/j.rehab.2009.04.003. PubMed PMID: 19783494.
18. Ramesh AN, Kambhampati C, Monson JR, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl*. 2004;86(5):334-8. Epub 2004/08/31. doi: 10.1308/147870804290. PubMed PMID: 15333167; PubMed Central PMCID: PMCPMC1964229.
19. Awad M, Khanna R. *Machine Learning. Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*. Berkeley, CA: Apress; 2015. p. 1-18.
20. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med Inform Decis Mak*. 2019;19(1):281. Epub 2019/12/23. doi: 10.1186/s12911-019-1004-8. PubMed PMID: 31864346; PubMed Central PMCID: PMCPMC6925840.
21. Biswas M, Kuppili V, Saba L, Edla DR, Suri HS, Cuadrado-Godia E, et al. State-of-the-art review on deep learning in medical imaging. *Front Biosci (Landmark Ed)*. 2019;24:392-426. Epub 2018/11/24. PubMed PMID: 30468663.
22. Kohli M, Prevedello LM, Filice RW, Geis JR. Implementing Machine Learning in Radiology Practice and Research. *AJR Am J Roentgenol*. 2017;208(4):754-60. Epub 2017/01/27. doi: 10.2214/ajr.16.17224. PubMed PMID: 28125274.
23. Jin R, Luk KD, Cheung JPY, Hu Y. Prognosis of cervical myelopathy based on diffusion tensor imaging with artificial intelligence methods. *NMR Biomed*. 2019;32(8):e4114. Epub 2019/05/28. doi: 10.1002/nbm.4114. PubMed PMID: 31131933.
24. Tack C. Artificial intelligence and machine learning | applications in musculoskeletal physiotherapy. *Musculoskelet Sci Pract*. 2019;39:164-9. Epub 2018/12/07. doi: 10.1016/j.msksp.2018.11.012. PubMed PMID: 30502096.
25. Vernon H. The Neck Disability Index: state-of-the-art, 1991-2008. *J Manipulative Physiol Ther*. 2008;31(7):491-502. Epub 2008/09/23. doi: 10.1016/j.jmpt.2008.08.006. PubMed PMID: 18803999.
26. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil*. 2008;89(1):69-74. Epub 2008/01/01. doi: 10.1016/j.apmr.2007.08.126. PubMed PMID: 18164333.
27. Meisingset I, Stensdotter AK, Woodhouse A, Vasseljen O. Neck motion, motor control, pain and disability: A longitudinal study of associations in neck pain patients in physiotherapy treatment. *Man Ther*. 2016;22:94-100. Epub 2015/11/21. doi: 10.1016/j.math.2015.10.013. PubMed PMID: 26586133.
28. Meisingset I, Woodhouse A, Stensdotter AK, Stavdahl O, Loras H, Gismervik S, et al. Evidence for a general stiffening motor control pattern in neck pain: a cross sectional study. *BMC Musculoskelet Disord*. 2015;16:56. Epub 2015/04/19. doi: 10.1186/s12891-015-0517-2. PubMed PMID: 25888215; PubMed Central PMCID: PMCPMC4377005.
29. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain*. 2004;8(4):283-91. Epub 2004/06/23. doi: 10.1016/j.ejpain.2003.09.004. PubMed PMID: 15207508.
30. Boonstra AM, Stewart RE, Köke AJ, Oosterwijk RF, Swaan JL, Schreurs KM, et al. Cut-Off Points for Mild, Moderate, and Severe Pain on the Numeric Rating Scale for Pain in Patients with Chronic Musculoskeletal Pain: Variability and Influence of Sex and Catastrophizing. *Front Psychol*. 2016;7:1466. Epub 2016/10/18. doi: 10.3389/fpsyg.2016.01466. PubMed PMID: 27746750; PubMed Central PMCID: PMCPMC5043012.

31. Schneider GM, Jull G, Thomas K, Smith A, Emery C, Faris P, et al. Derivation of a clinical decision guide in the diagnosis of cervical facet joint pain. *Arch Phys Med Rehabil.* 2014;95(9):1695-701. Epub 2014/03/26. doi: 10.1016/j.apmr.2014.02.026. PubMed PMID: 24662813.
32. Hage R, Detrembleur C, Dierick F, Pitance L, Jojczyk L, Estievenart W, et al. DYSKIMOT: An Ultra-Low-Cost Inertial Sensor to Assess Head's Rotational Kinematics in Adults during the Didren-Laser Test. *Sensors (Basel).* 2020;20(3). Epub 2020/02/09. doi: 10.3390/s20030833. PubMed PMID: 32033169.
33. Kononenko I. Machine learning for medical diagnosis: history, state of the art and perspective. *Artif Intell Med.* 2001;23(1):89-109. Epub 2001/07/27. doi: 10.1016/s0933-3657(01)00077-x. PubMed PMID: 11470218.
34. Berrar D. Cross-Validation. *Encyclopedia of Bioinformatics and Computational Biology.* 2018;1:542-5. doi: 10.1016/B978-0-12-809633-8.20349-X.
35. Ndiaye E LT, . Fercoq O, . Salmon J, . Takeuchi I, . Safe grid search with optimal complexity. *Proceeding of the 36th International Conference on Machine Learning.* 2019.
36. Kumar VM, Sonajharia. Feature Selection: A literature Review. *The Smart Computing Review.* 2014;4(3). doi: 10.6029/smartcr.2014.03.007.
37. Franov E, Straub M, Bauer CM, Ernst MJ. Head kinematics in patients with neck pain compared to asymptomatic controls: a systematic review. *BMC Musculoskelet Disord.* 2022;23(1):156. Epub 2022/02/18. doi: 10.1186/s12891-022-05097-z. PubMed PMID: 35172799.
38. Fukuchi RK, Eskofier BM, Duarte M, Ferber R. Support vector machines for detecting age-related changes in running kinematics. *J Biomech.* 2011;44(3):540-2. Epub 2010/10/29. doi: 10.1016/j.jbiomech.2010.09.031. PubMed PMID: 20980005.
39. Lai DT, Levinger P, Begg RK, Gilleard WL, Palaniswami M. Automatic recognition of gait patterns exhibiting patellofemoral pain syndrome using a support vector machine approach. *IEEE Trans Inf Technol Biomed.* 2009;13(5):810-7. Epub 2009/05/19. doi: 10.1109/TITB.2009.2022927. PubMed PMID: 19447723.
40. Suda EY, Watari R, Matias AB, Sacco ICN. Recognition of Foot-Ankle Movement Patterns in Long-Distance Runners With Different Experience Levels Using Support Vector Machines. *Front Bioeng Biotechnol.* 2020;8:576. Epub 2020/07/01. doi: 10.3389/fbioe.2020.00576. PubMed PMID: 32596226; PubMed Central PMCID: PMC7300177.
41. Jiang Y, Sadeqi A, Miller EL, Sonkusale S. Head motion classification using thread-based sensor and machine learning algorithm. *Sci Rep.* 2021;11(1):2646. Epub 2021/01/31. doi: 10.1038/s41598-021-81284-7. PubMed PMID: 33514762; PubMed Central PMCID: PMC7846730.
42. Jairam V, Park HS. Strengths and limitations of large databases in lung cancer radiation oncology research. *Transl Lung Cancer Res.* 2019;8(Suppl 2):S172-S83. Epub 2019/11/02. doi: 10.21037/tlcr.2019.05.06. PubMed PMID: 31673522; PubMed Central PMCID: PMC6795574.
43. Harper DE, Shah Y, Ichescio E, Gerstner GE, Peltier SJ. Multivariate classification of pain-evoked brain activity in temporomandibular disorder. *Pain Rep.* 2016;1(3). Epub 2017/06/21. doi: 10.1097/pr9.0000000000000572. PubMed PMID: 28630949; PubMed Central PMCID: PMC5473632.
44. Guo R, Zhou C, Wang C, Tsai TY, Yu Y, Wang W, et al. In vivo primary and coupled segmental motions of the healthy female head-neck complex during dynamic head axial rotation. *J Biomech.* 2021;123:110513. Epub 2021/05/27. doi: 10.1016/j.jbiomech.2021.110513. PubMed PMID: 34038861; PubMed Central PMCID: PMC8227877.
45. Moghaddas D, de Zoete RMJ, Edwards S, Snodgrass SJ. Differences in the kinematics of the cervical and thoracic spine during functional movement in individuals with or without chronic neck

- pain: a systematic review. *Physiotherapy*. 2019;105(4):421-33. Epub 2019/04/22. doi: 10.1016/j.physio.2019.01.007. PubMed PMID: 31005251.
46. Betsch MW, Blizzard SR, Shinseki MS, Yoo JU. Prevalence of degenerative changes of the atlanto-axial joints. *Spine J*. 2015;15(2):275-80. Epub 2014/10/04. doi: 10.1016/j.spinee.2014.09.011. PubMed PMID: 25277533.
47. Pan F, Arshad R, Zander T, Reitmaier S, Schroll A, Schmidt H. The effect of age and sex on the cervical range of motion - A systematic review and meta-analysis. *J Biomech*. 2018;75:13-27. Epub 2018/05/20. doi: 10.1016/j.jbiomech.2018.04.047. PubMed PMID: 29776822.
48. Liew BXW, Rugamer D, De Nunzio AM, Falla D. Interpretable machine learning models for classifying low back pain status using functional physiological variables. *Eur Spine J*. 2020;29(8):1845-59. Epub 2020/03/04. doi: 10.1007/s00586-020-06356-0. PubMed PMID: 32124044.
49. Falla D, Devecchi V, Jiménez-Grande D, Rügamer D, Liew BXW. Machine learning approaches applied in spinal pain research. *J Electromyogr Kinesiol*. 2021;61:102599. Epub 2021/10/09. doi: 10.1016/j.jelekin.2021.102599. PubMed PMID: 34624604.