

Automatic selection model to identify neurodegenerative diseases

DIGITAL HEALTH
Volume 10: 1–11
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076241284376
journals.sagepub.com/home/dhj



Eddy Sánchez-DelaCruz¹ , Cecilia-Irene Loeza-Mejía¹ ,
César Primero-Huerta^{1,2} and Mirta Fuentes-Ramos¹

Abstract

Objective: This study evaluates machine learning algorithms' effectiveness in classifying Parkinson's disease and Huntington's disease based on biomarker data obtained non-invasively from patients and healthy controls.

Methods: Datasets containing biomarker data (x , y , and z values of accelerometers) from sensors were collected from Parkinson's disease, Huntington's disease patients, and healthy controls. An automatic selection model method was implemented for disease classification, using a unique Mexican database of human gait biomarkers, which we consider the only one of its kind. Random forest, random subspace method, and K-star algorithms were employed, with parameters optimized through an automated model selection.

Results: The study achieved a 0.893 precision rate for Parkinson's disease and Huntington's disease using the random subspace method. The findings underscore the potential of machine learning techniques in medical diagnosis, particularly in neurological disorders.

Conclusion: The automatic selection model method demonstrated efficacy in classifying Parkinson's disease and Huntington's disease based on non-invasive biomarker data. This research contributes to advancing non-invasive diagnostic approaches in neurological disorders, highlighting the significance of machine learning in healthcare.

Keywords

Parkinson's disease, Huntington's disease, machine learning, biomarkers, automated machine learning

Submission date: 6 March 2024; Acceptance date: 28 August 2024

Introduction

The nervous system comprises the central nervous system (CNS) and the peripheral nervous system.¹ These systems oversee numerous functions in the human body, and various conditions can disturb their homeostasis. Notably, neurodegenerative disorders, as discussed in previous research, are characterized by progressive deterioration that impacts fundamental abilities such as movement and cognition.^{2,3} Both Parkinson's disease (PD) and Huntington's disease (HD) are part of this group and greatly affect the patient's quality of life. However, due to their similar early symptoms, distinguishing between them can be challenging. In addition, for their diagnosis techniques that are highly expensive and sometimes invasive for the patient are

used,⁴ so it is important to find alternatives for their classification.

PD is a neurological disorder marked by autonomic and motor disturbances.⁵ It is estimated that 0.3% of the population in industrialized countries is affected by this

¹Artificial Intelligence Laboratory, Tecnológico Nacional de México/Instituto Tecnológico Superior de Misantla, Veracruz, Mexico

²División de Ingeniería en Sistemas Computacionales, Tecnológico Nacional de México/Valle de Bravo, Valle de Bravo Mexico

Corresponding author:

Sánchez-DelaCruz, Artificial Intelligence Laboratory, Tecnológico Nacional de México/Instituto Tecnológico Superior de Misantla, Km. 1.8 Carretera a Loma del Cojolite, Misantla, Veracruz 93821, Mexico.
Email: eddsax@gmail.com



condition, with men being twice as likely as women to develop the disease. Additionally, the risk of developing PD increases with age.^{6,7}

This condition arises from the degeneration of dopaminergic neurons in brain areas such as the substantia nigra, resulting in decreased levels of this neurotransmitter in the striatum,⁸ as well as possible degeneration of catecholaminergic and serotonergic neurons, causing altered movement patterns.⁹

On the other hand, HD is also a neurological disorder of the CNS, whose clinical manifestations are chorea, involuntary movements such as bradykinesia, loss of postural reflexes, ataxia, and gait, with symptoms appearing between 15 and 20 years of age.^{11,10} It is known that 10–13 persons are prevalent per 100,000 in America, being more significant in women than men. The symptoms are more aggressive in women due to differences in the sequences of genes that code for the protein huntingtin, which causes the disease.^{12,13}

Likewise, it is known that the disease is due to alterations in the gene sequence in chromosome 4, impacting brain structures like the striatum, and leading to atrophy of the cerebral cortex and cerebellum, which manifests as movement disorders.¹⁴

Currently, there are studies that due to the similarity of the symptoms try to differentiate these diseases through biomarkers or the use of techniques such as machine learning (ML), using algorithms to databases provided by different types of studies.¹⁵ Therefore, our contributions are (1) implementing a method based on an automatic selection model to identify (classify) neurodegenerative diseases such as PD and HD, and (2) utilizing the database with human gait biomarkers obtained from the National Institute of Neurology and Neurosurgery (NINN), which we believe is the only database on this topic. The article will first describe the previous works that support the research, continuing with the materials and methods used, that is, the sensor network, the dataset, the ML model, and the metrics, then the results, a detailed discussion of the results, and the conclusions with future work planned to enhance the study.

Previous works

Currently, there exists a diverse array of research that has enabled the classification of neurodegenerative conditions such as PD, HD, ataxia, amyotrophic lateral sclerosis, and spinal muscular atrophy employing various biological as well as motor markers;^{16,17} in this regard, those that use gait as a differentiator of PD and HD will be described.

Gait and ML as tools to differentiate neurodegenerative diseases: PD and HD

It is known that in PD, there are alterations in both movement and posture; balance, hip position, and gait, the

latter has functioned as an indication of the progress of the disease,¹⁸ through the measurement of kinematic and spatiotemporal parameters, by the use of 3D software, however, some bias may occur because they are observational studies.^{19,20} On the other hand, there are also studies where using sensors in lower extremities (soles of the feet), databases of patients with idiopathic hyposmia, PD, and healthy subjects have been obtained, resulting in high accuracy values of 97% through the random forest algorithm.²¹ Likewise, using inertial measurement units in PD patients and healthy patients in the lower extremities and back capturing acceleration and rotational motion, high classification efficiency values have been achieved using support vector machine (SVM) algorithm.³ And by using wireless inertial sensors that measure head, pitch, roll, and stride rotations and analysis with ML techniques such as SVM we can correctly classify subjects with PD and healthy subjects with values above 90%.²² In addition, one of the most important in our opinion has been the use of smartphones (due to their low cost) and their sensors (gyroscope) to know the severity of PD through developed applications and disease severity score learning (DSSL) algorithms with efficiency values higher than 90%.²³

Regarding HD, it is also known that one of the main characteristics is gait impairment, serving as an element for its classification.^{25,24,26} In this way, there are studies where it has been possible to differentiate the disease and its progression using algorithms.²⁷ For instance, a 2016 study utilized inertial sensors to assess elderly individuals post-stroke and HD patients, achieving SVM algorithm classification accuracy of 90.5%, relative to healthy subjects.²⁸ Similarly, using the Unified Huntington's Disease Rating Scale, using biometric sensors on the trunk and extremities (triaxial accelerometer and gyroscope) or foot pressure sensors, the linear discriminant analysis (LDA) and VGG16 algorithms have been reported as the most optimal for measuring disease severity, obtaining percentages of 96% and 89%, respectively, comparing healthy subjects and those with HD.^{29,30}

Analysis and opportunity gap

All these studies have been very important and relevant since ML is really strengthening the patient's diagnosis in contribution to the medical staff. Gait is a fundamental parameter in the evaluation of neurodegenerative diseases,^{32,31,33,28,24} as it provides important information about the patient's motor function and overall neurological health. Sensors that monitor gait through inertial signals, such as those obtained from accelerometers and gyroscopes, can be of great help in the accurate evaluation and early detection of such conditions.^{31,3,28,21,25} At present, there are very few studies that employ these techniques using gait as a differentiator and fewer in Latin American countries. Therefore, the present work represents great advances in medicine in

Mexico. Table 1 shows some examples of analyses using ML in some neurodegenerative diseases and the countries in which the study has been carried out.

Materials and methods

The development of the project went from obtaining the dataset by NINN using body sensors, and processing the data in Weka software (as well as searching for algorithms), to algorithm selection and acquisition of selection values (see Figure 1).

Body sensor network

The network of three-axis ADXL-335 accelerometer sensors used to obtain the measurements and connected to an Arduino MEGA-2560 were placed on the extremities of both knees, ankles, and thorax, analyzing the Cartesian x , y , and z axes. From this network of sensors in Figure 2, which are accessible to be acquired by the patient, the dataset was obtained from Fuentes-Ramos et al.¹⁶

Dataset

Data for the study were collected and ethically approved by the NINN Ethics Committee.^{40,16} Gait biomarkers were employed in the evaluation of patients with PD and HD, along with a control group comprising healthy subjects.

In addition, patients and family members signed informed consent forms to ensure that the data were published and that they did not reveal the identity of the participants.

The total number of patients was 78, which were divided into 47 with PD, 13 with HD, and 19 healthy subjects, 34 women and 45 men (see Table 2).

During the gait analysis process, data were provided in ~ 2 minutes.

The exclusion criteria were people who had difficulty walking by themselves or who were in wheelchairs.

The datasets comprising raw data without any preprocessing, were grouped in a single file for each participant and the information was united in a single document, both for those with the disease and those who were healthy. This

Table 1. Neurodegenerative diseases where ML has been used.

Disease	Country/region	Algorithm used	Key findings
Alzheimer	USA, Iran, Portugal	SVM. ^{32,34,31}	Gait data and ML can serve as objective tools for the early detection of cognitive impairment.
Amyotrophic lateral sclerosis	Republic of China, Taiwan	SVM, ³³ K-nearest neighbors. ³⁵	Gait variability can diagnose and monitor amyotrophic lateral sclerosis.
Ataxia	China, Czech Republic, Italy	Logistic regression, linear SVM, poly SVM, RBF SVM, Naïve Bayes, nearest neighbors, decision tree, random forest, neural net, AdaBoost, and multiplayer perceptron. ^{36,38,37}	Tools like kinect and ML algorithms are effective for assessing and classifying the severity of neurodegenerative diseases and ataxias.
Dementia	USA, Cuba, UK	Poisson regression analyses. ³⁹	Neurological impairments and gait disturbances are associated with dementia and mortality.
Parkinson's disease	Australia, USA, England, Italy	Random forest, ²¹ SVM, ^{3,21} DSSL, ²³ Naïve Bayes ²¹	High accuracy (97%) using sensors in lower extremities. ²¹ SVM achieved classification efficiency above 90% using inertial measurement units. ³ Smartphone sensors (gyroscope) with DSSL algorithms showed efficiency values above 90%. ²³
Huntington's disease	Italy, USA, Australia, Netherlands	SVM, ²⁸ LDA, ²⁹ VGG16 ³⁰	SVM achieved accuracies above 90% using inertial sensors. ²⁸ LDA and VGG16 reported accuracies of 96.4% and 89%, respectively, using biometric sensors. ^{29,30} Gait impairment used as a classification element. ^{25,24,26}

ML: machine learning; SVM: support vector machine; DSSL: disease severity score learning; LDA: linear discriminant analysis; RBF: radial basis function.

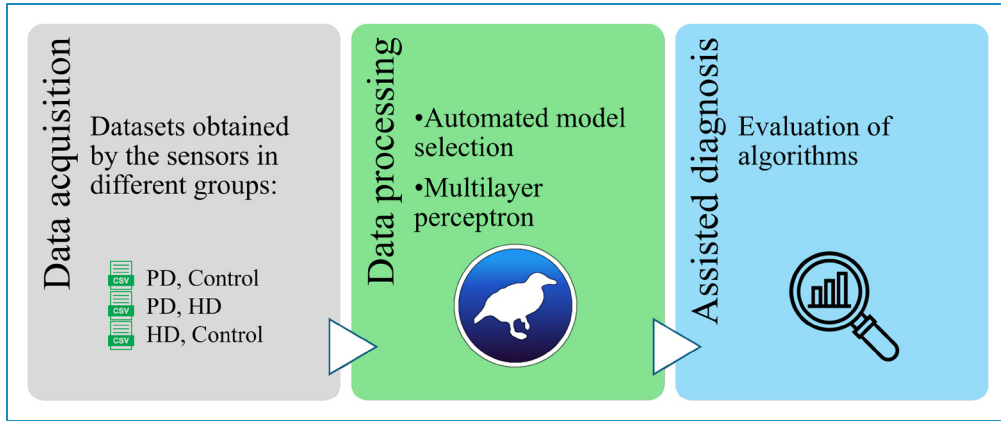


Figure 1. Pipeline of PD and HD categorization by using machine learning. PD: Parkinson’s disease; HD: Huntington’s disease.

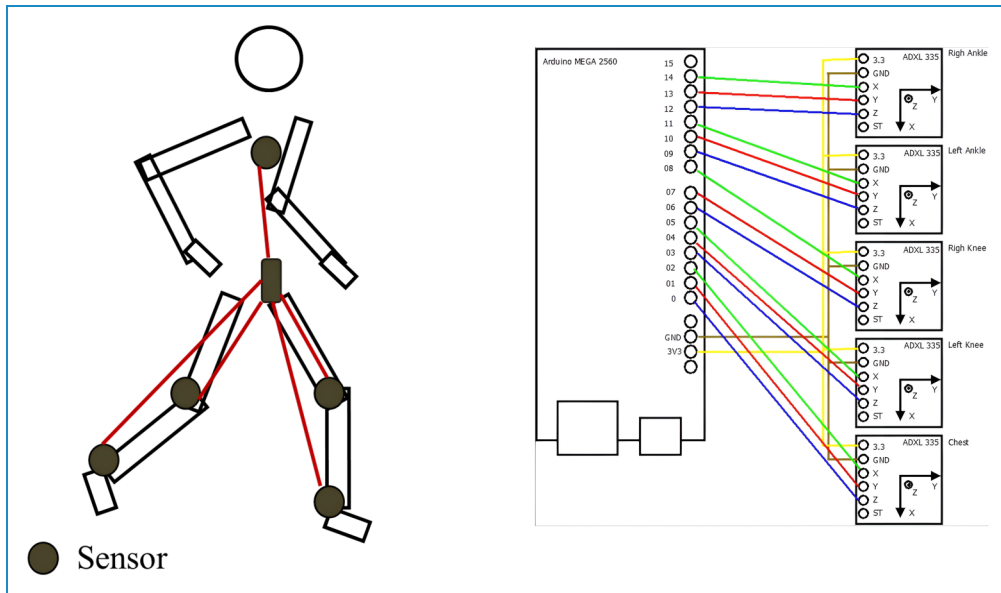


Figure 2. Examples of the sensors used to obtain the dataset.¹⁶

Table 2. Sex and age distribution of patients under study.

Case	Age groups			Sex	
	18-39	40-59	60-84	F	M
Control	5	9	5	12	7
PD	2	14	31	17	30
HD	5	7	1	5	8

PD: Parkinson’s disease; HD: Huntington’s disease.

resulted in three datasets, which were used to carry out the classification, resulting in a binary category, Each of these

datasets contains 1800 records per class, accumulating a total of 3600 records:

- Binary sets: {PD, Control}, {PD, HD} and {HD, Control}

Finally, a subject-wise strategy was employed to split the data, allocating 80% for training and the remaining 20% for testing.

Automatic selection model

Thornton et al.’s ML model selection method was employed to automatically choose the classification algorithm,^{42,41} which involves:

Given a collection of algorithms \mathcal{A} and a finite dataset $\mathcal{D} = (x_1, y_1), \dots, (x_n, y_n)$ for training, the objective is to

identify the algorithm A from \mathcal{A} demonstrates the best generalization capabilities. This process involves splitting \mathcal{D} into several training subsets, denoted as $\mathcal{D}^{(i)}_{train}$, and corresponding non-overlapping validation subsets, denoted as $\mathcal{D}^{(i)}_{valid}$ for $i = 1, \dots, k$. The learning function f_i is then derived by applying A to $\mathcal{D}^{(i)}_{train}$, and the effectiveness of these functions is evaluated on $\mathcal{D}^{(i)}_{valid}$. This approach captures the challenge of selecting the most effective algorithm for the given task.

$$A^* \in [A \in \mathcal{A}] \operatorname{argmin} \frac{1}{k} \cdot \sum_{i=1}^k \mathcal{L}(A, \mathcal{D}^{(i)}_{train}, \mathcal{D}^{(i)}_{valid}) \quad (1)$$

Here, $\mathcal{L}(A, \mathcal{D}^{(i)}_{train}, \mathcal{D}^{(i)}_{valid})$ represents the loss, specifically the misclassification rate, incurred by algorithm A during training on $\mathcal{D}^{(i)}_{train}$ and its subsequent evaluation on $\mathcal{D}^{(i)}_{valid}$. The cross-validation technique is employed to partition the training data into k sets of equal size, denoted as $\mathcal{D}^{(1)}_{valid}, \dots, \mathcal{D}^{(k)}_{valid}$, with the complementary training sets defined as $\mathcal{D}^{(i)}_{train} = \mathcal{D} \setminus \mathcal{D}^{(i)}_{valid}$ for $i = 1, \dots, k$.

Based on the above equation and the experimental tests performed in Waikato Environment Knowledge Analysis v.3.8,⁴³ it was found that the random forest algorithms for the set {PD, Control} were the most appropriate algorithms for classification. It is important to point out that the analysis was performed using an exhaustive algorithm-by-algorithm search provided by the software. The mathematical basis of the selected algorithms will be shown below.

Algorithms

Random forest. This algorithm, as a key component of the methods employed, utilizes an ensemble of classifiers denoted as $h(x, \Theta_k)$, $k = 1, \dots$, where Θ_k represents independent random vectors.⁴⁴ The algorithm is outlined below:

Algorithm 1 Random forest.^{16,44}

Input : dataset $T = (x, y)$, number of trees m , number of random levels k

Output : RF, a set of grown trees

Initialization RF

for $i = 1$ to m **do**

 Create a bootstrap sample T' from dataset T

 Grow a decision tree Tree using T' and parameter k

 Add Tree to RF

end

This technique includes the inputs which are the data (T) containing previous information (x) with its result (y), the number of decision trees (m), as well as the highest value that each tree will have (k). Once the values are obtained, new training data (T') are elaborated with sampling with replacement. Randomly selected data is processed using two functions, choosing the most appropriate partition. This is done repeatedly until all features have been analyzed.⁴⁴

K-star. The definition of K^{*45} established by examining a set \mathbf{I} of instances, potentially infinite, and a finite collection of transformations \mathbf{T} applied to \mathbf{I} . Each transformation $t \in \mathbf{T}$ is a function that maps instances to other instances, represented as $t: \mathbf{I} \rightarrow \mathbf{I}$. Within \mathbf{T} there exists a specific element σ (known as the stop symbol) that, for completeness, maps instances to themselves ($\sigma(a) = a$). Consider \mathbf{P} as the set of all prefix codes generated from \mathbf{T}^* and terminated by σ . Members of \mathbf{T}^* and consequently \mathbf{P} uniquely define transformations on \mathbf{I} :

$$\bar{t}(a) = t_n(t_{n-1}(\dots t_1(a)\dots)), \text{ where } \bar{t} = t_1, \dots, t_n \quad (2)$$

A probability function p is established on \mathbf{T}^* , adhering to the following properties:

$$\begin{aligned} 0 &\leq \frac{p(\bar{t}u)}{p(\bar{t})} \leq 1 \\ \sum_u p(\bar{t}u) &= p(\bar{t}) \\ p(\Lambda) &= 1 \end{aligned} \quad (3)$$

As a result, it adheres to the following equation:

$$\sum_{\bar{t} \in \mathbf{P}} p(\bar{t}) = 1 \quad (4)$$

The probability function P^* is established to represent the likelihood of all paths connecting instance a to instance b :

$$P^*(b | a) = \sum_{\bar{t} \in \mathbf{P}: \bar{t}(a)=b} p(\bar{t}) \quad (5)$$

It can be readily demonstrated that P^* adheres to the following properties:

$$\begin{aligned} \sum_b P^*(b | a) &= 1 \\ 0 &\leq P^*(b | a) \leq 1 \end{aligned} \quad (6)$$

The K^* function is then defined as follows:

$$K^*(b | a) = -\log_2 P^*(b | a) \quad (7)$$

K^* does not strictly adhere to the typical traits of a distance function. For instance, $K^*(aaa)$ is generally non-zero, and the function, as denoted by the I notation, lacks symmetry. Despite the potential counter-intuitiveness of these properties, their absence does not impede the progression of the K^* outlined below. The following properties can be demonstrated:

$$\begin{aligned} K^*(b | a) &\geq 0 \\ K^*(c | b) + K^*(b | a) &\geq K^*(c | a) \end{aligned} \quad (8)$$

Random subspace method. This method constructs a decision forest ensemble,⁴⁶ using S training samples represented by X_j p -dimensional vectors, where $X_j = (x_{j1}, x_{j2}, \dots, x_{jp})$. This method automatically selects p^* features with $p^* < p$.⁴⁷ The random subspace algorithm is shown in Algorithm 2.

Algorithm 2 Random subspace method.^{46,47}

Input : Training set S , total of subspaces B , subspace dimension p^*

Output : Ensemble E

Initialize E to empty set

for $i = 1$ to B do

$\tilde{S}^i \leftarrow \text{SelectRandomSubspace}(\tilde{S}, p^*)$

$C^i \leftarrow \text{ConstructClassifier}(\tilde{S}^i)$

$E \leftarrow E \cup \{C^i\}$

end

Multilayer perceptron. This type of artificial neural network consists of an input layer, one or more hidden layers, and an output layer.⁴⁸ Each layer contains multiple neurons, and every neuron in one layer is connected to all neurons in the next layer. This architecture enables the multilayer perceptron to model complex, non-linear relationships in data. The input features are represented in

$$\mathbf{x} = (x_1, x_2, \dots, x_n) \quad (9)$$

In the hidden layers, each neuron computes a weighted sum of inputs,^{50,49} adds a bias, and then applies an activation function f :

$$h_j^{(l)} = f\left(\sum_i w_{ji}^{(l)} h_i^{(l-1)} + b_j^{(l)}\right) \quad (10)$$

where $h_i^{(l-1)}$ is the output from the previous layer, $w_{ji}^{(l)}$ are weights, and $b_j^{(l)}$ are biases. The output neurons compute:

$$o_j = g\left(\sum_i w_{ji}^{(L+1)} h_i^{(L)} + b_j^{(L+1)}\right) \quad (11)$$

where g is the activation function for the output layer.

Metrics

To assess performance, the confusion matrix (see Table 3), Kappa (equation (12)), precision (equation (13)), sensitivity or recall (equation (14)), f -measure (equation (15)), and area under the receiver operating characteristic curve were used as evaluation metrics. In a confusion matrix, the counts of predicted classes are displayed in columns, while actual

Table 3. Confusion matrix.

		True class	
		Positive	Negative
Predicted class	Positive	TP	FP
	Negative	FN	TN

TP: true positive; FP: false positive; FN: false negative; TN: true negative.

values are shown in rows. This matrix helps identify true negative (TN), true positive (TP), false negative (FN), and false positive (FP). The confusion matrix provides the TP rate of correctly classified instances. The fraction of instances classified in the positive class is obtained by precision. The F value integrates the characteristics of the PT rate, and the precision becomes a single factor. While the receiver operating characteristic curve illustrates the TP rate and the FP rate.

$$\kappa = \frac{2 \cdot (TP \cdot TN - FP \cdot FN)}{(TP + FP) \cdot (FP + TN) + (TP + FN) \cdot (FN + TN)} \quad (12)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (13)$$

$$\text{TP rate} = \frac{TP}{TP + FN} \quad (14)$$

$$F\text{-measure} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (15)$$

Experiments

The experimental setup involved the evaluation of two main cases for the classification of biomedical data into binary sets: the *first* using the automatic selection model and the *second* employing the multilayer perceptron. The classifications performed were:

- PD versus control: In the first case, the random forest algorithm was selected using the automatic selection model, while in the second case, the multilayer perceptron was employed.
- PD versus HD: The random subspace method was automatically selected in the first case, and the multilayer perceptron was used in the second.
- HD versus control: For this comparison, the K-star algorithm was automatically selected in the first case, and the multilayer perceptron was used in the second.

The performance metrics include the percentage of correctly classified instances, the Kappa statistic, weighted average precision, weighted average recall, and weighted average F -measure. Additionally, the confusion matrix for each classification task is provided.

The experiments were performed in Waikato Environment Knowledge Analysis v.3.8,⁴³ on an HP laptop with Windows 11 64-bit, Intel (R) Core(TM) i7-1065G7 processor @ 1.30 GHz, and 12.00 GB RAM.

Results

The classification results are presented in Table 4. In the *first case*, the automatic selection model identified the following three algorithms for the classification of the binary sets:

Table 4. Classification results of binary sets.

Subsets	Algorithm	% Correctly classified	Kappa statistic	Precision	Recall	F-measure	ROC area	Confusion matrix		
(PD,control)	Random	91.3889	0.8278	0.914	0.914	0.914	0.97	<i>a</i>	<i>b</i>	
	forest							Control = <i>a</i>	326	35
								PD = <i>b</i>	27	332
(PD, HD)	Random	89.3056	0.7861	0.893	0.893	0.893	0.786	<i>a</i>	<i>b</i>	
	subspace							HD = <i>a</i>	323	38
								PD = <i>b</i>	39	320
(HD,control)	K-star	80.5556	0.6111	0.806	0.806	0.806	0.611	<i>a</i>	<i>b</i>	
								Control = <i>a</i>	285	76
								HD = <i>b</i>	64	295
(PD,control)	Multilayer	83.4722	0.6695	0.835	0.835	0.835	0.894	<i>a</i>	<i>b</i>	
	perceptron							Control = <i>a</i>	295	66
								PD = <i>b</i>	53	306
(PD, HD)	Multilayer	77.9167	0.5583	0.779	0.779	0.779	0.815	<i>a</i>	<i>b</i>	
	perceptron							HD = <i>a</i>	284	77
								PD = <i>b</i>	82	277
(HD,control)	Multilayer	64.5833	0.2915	0.646	0.646	0.646	0.696	<i>a</i>	<i>b</i>	
	perceptron							Control = <i>a</i>	246	115
								HD = <i>b</i>	140	219

PD: Parkinson's disease; HD: Huntington's disease; ROC: receiver operating characteristic.

- PD versus control: The random forest algorithm correctly identified 91.3889% of the cases with a Kappa statistic of 0.8278, and precision, recall, and *F*-measure all at 0.914. The confusion matrix indicates that 326 control instances and 332 PD instances were correctly classified.
 - PD versus HD: Using the random subspace method, 89.3056% of the cases were correctly identified with a Kappa statistic of 0.7861. Precision, recall, and *F*-measure were all 0.893. The confusion matrix shows that 323 HD instances and 320 PD instances were correctly classified.
 - HD versus control: The K-star algorithm correctly identified 80.5556% of the cases with a Kappa statistic of 0.6111. Precision, recall, and *F*-measure were all 0.806. The confusion matrix highlights that 285 control instances and 295 HD instances were correctly classified.
- In the second case, a multilayer perceptron was used for the classification of the binary datasets, yielding the following results:
- PD versus control: The multilayer perceptron correctly identified 83.47% of the data with a Kappa statistic of 0.6695. Precision, recall, and *F*-measure were all 0.835. The confusion matrix shows that 295 instances of control and 306 instances of PD were classified correctly.
 - PD versus HD: Using the multilayer perceptron, 77.92% of the data was correctly identified with a Kappa statistic of 0.5583. Precision, recall, and *F*-measure were all 0.779. The confusion matrix shows 284 instances of HD and 277 instances of PD classified correctly.

- HD versus control: The multilayer perceptron correctly identified 64.58% of the data with a Kappa statistic of 0.2915. Precision, recall, and F -measure were all 0.646. The confusion matrix highlights 246 instances of control and 219 instances of HD classified correctly.

Figures 3 to 5 display the precision–recall graphs, demonstrating values around 85%.

Discussion

In the present study, we found some interesting findings. For example, it is the first study that presents a comparison on automatic selection model and multilayer perceptron on a dataset of patients in Mexico (NINN), so it represents a valuable tool for physicians in this country because it has been documented, for example, that the genetic etiology and prevalence of both (PD and HD) and therefore the characteristics of involuntary movements are different in different sites in Latin America and the world, so using this technique is vital at present.^{53,51,52}

Likewise, the sensor network used in the study subjects is unique, since it has been designed in such a way that the data obtained in the different X , Y , and Z axes of the different limbs provide greater information for the dataset.¹⁶

The results of the experiments show significant differences in the performance of the algorithms used in this study, highlighting the importance of proper algorithm selection for different biomedical data classification tasks.

In the first case, the automatic selection model identified three algorithms: random forest, random subspace, and K-star. These algorithms exhibited high performance in terms of precision and Kappa statistic. Specifically,

random forest achieved the highest precision (0.914) and the highest Kappa value (0.8278) in classifying PD versus control, indicating excellent ability of the model to distinguish between these two classes.

The random subspace method also yielded good results in classifying between PD and HD, with a precision of 0.893 and a Kappa statistic of 0.7861. Although this algorithm showed slightly lower performance compared to Random forest, it still achieved accurate classification.

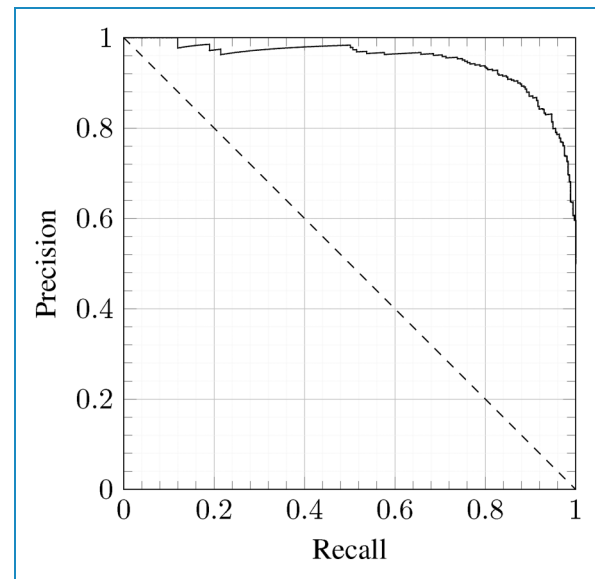


Figure 4. Precision and recall graph of PD versus HD using random subspace. PD: Parkinson's disease; HD: Huntington's disease.

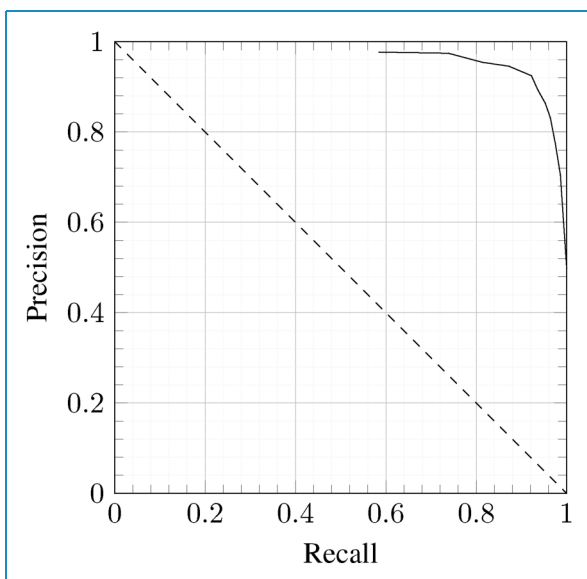


Figure 3. Precision and recall graph of Parkinson's disease (PD) versus control using random forest.

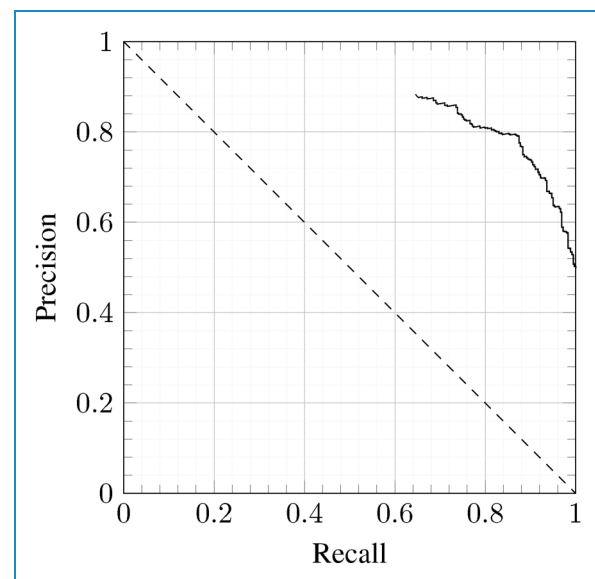


Figure 5. Precision and recall graph for Huntington's disease (HD) versus control using K-star.

The K-star algorithm, while performing acceptably with a precision of 0.806 and a Kappa statistic of 0.6111 in classifying between HD and control, was the least effective among the three selected by the automatic selection model.

In the second case, the multilayer perceptron was used for all classification tasks. Although its performance was inferior compared to the algorithms selected automatically, the multilayer perceptron still showed reasonable results. The highest precision was achieved in the classification between PD and Control (0.835) with a Kappa statistic of 0.6695. However, the confusion matrix reveals a higher number of misclassifications compared to random forest.

For the classification between PD and HD, the multilayer perceptron achieved a precision of 0.779 and a Kappa statistic of 0.5583. These results are lower compared to those obtained with the random subspace method, suggesting that the multilayer perceptron may not be the most suitable model for this specific task.

Finally, in the classification between HD and Control, the multilayer perceptron showed the lowest performance with a precision of 0.646 and a Kappa statistic of 0.2915. The confusion matrix indicates more difficulties in distinguishing between these two classes, with a significant number of misclassifications.

The results suggest that automatically selected classification algorithms tend to outperform the multilayer perceptron in terms of precision and capacity to discriminate between classes. This underscores the importance of using automated model selection to identify the most suitable algorithms for specific classification tasks in complex datasets such as biomedical data. Additionally, it is noteworthy that for both approaches analyzed, the best performance was achieved in the classification between PD and control, indicating a greater ability to distinguish between these two groups. In contrast, the classification between HD and control showed the poorest performance, highlighting potential additional challenges in differentiating these specific groups. This pattern suggests that the evaluated algorithms could be more effective in detecting subtle differences in closer groups in terms of clinical and pathological characteristics.

It is important to highlight that the analysis conducted in the Weka software involved a comprehensive search of algorithms, evaluating the dataset one by one from the package of options offered by the program. This approach differs from other studies where only specific algorithms or Autoweka (an algorithm with parameter optimization) are used, which may provide a more thorough exploration of possibilities.

Conclusions and perspectives

In conclusion, this study highlights the effectiveness of automated model selection in identifying high-performance algorithms with percentages close to 90% in the classification identification of PD and HD based on gait biomarkers. However, limitations of the multilayer perceptron are evident

in certain classification scenarios. These findings underscore the importance of customizing classification models to optimize performance in specific biomedical applications.

Furthermore, the variability in performance among different algorithms suggests that there is no single optimal model for all classification tasks.

Something very important is that the database used is from patients in different stages of the disease, so future works are expected to be carried out with data in the early stages of the disease. Therefore, the following studies are expected to be carried out:

- Developing and evaluating combinations of algorithms and advanced automatic selection techniques to further improve the accuracy and robustness of classification systems.
- Employ other types of devices, for example, smart bands, which have become very accessible to the patient and thus costs remain low.
- To use the algorithms found with a greater percentage of efficacy in other neurodegenerative pathologies that also affect gait, such as ataxias or amyotrophic lateral sclerosis, in both binary and multifactorial studies.
- To carry out ML studies in datasets of other chronic degenerative diseases, such as diabetes, obesity, hypertension, arthritis, asthma, etc., using clinical laboratory tests.

Acknowledgements: We would like to thank Dr Marie-Catherine Boll-Woehrlen, a specialist in movement disorders, for her advice and support with the tests carried out on the volunteers in the gait laboratory of the National Institute of Neurology and Neurosurgery from Mexico. The authors express their sincere gratitude to all the patients who participated in this study for their invaluable contribution and cooperation.

Contributorship: Eddy Sánchez-DelaCruz: project administration, supervision, methodology, formal analysis, and writing—original draft preparation. Cecilia-Irene Loeza-Mejía: conceptualization of this study, data curation, methodology, and writing—original draft preparation. César Primero-Huerta: investigation, methodology, and writing—original draft preparation. Mirta Fuentes-Ramos: conceptualization of this study, investigation, methodology, and writing—original draft preparation.

Declaration of conflicting interests: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: Ethical approval for this study was obtained from the National Institute of Neurology and Neurosurgery in Mexico, CONAHCyT project number FOMIX-TAB:2014-C29-245876. Also, a signed informed consent form was obtained from each patient and their family member prior to study initiation.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication

of this article: The authors would like to thank CONAHCyT, Mexico, for the fund 2021-000018-02NACF-12228, for graduate studies assigned to Loeza-Mejía.

Guarantor: National Institute of Neurology and Neurosurgery from Mexico.

ORCID iDs: Eddy Sánchez-DelaCruz  <https://orcid.org/0000-0002-2357-7799>

Cecilia-Irene Loeza-Mejía  <https://orcid.org/0000-0002-4584-0076>

References

- Sousa A, Meyer KA, Santpere G, et al. Evolution of the human nervous system function, structure and development. *Cell* 2017; 170: 226–247.
- Dugger BN and Dickson DW. Pathology of neurodegenerative diseases. *Cold Spring Harb Perspect Biol* 2017; 9: a028035.
- Ireland D, Wang Z, Lamont R, et al. Classification of movement of people with Parkinsons disease using wearable inertial movement units and machine learning. In: *Digital health innovation for consumers, clinicians, connectivity and community*. IOS Press, 2016, pp.61–66.
- Ross CA and Tabrizi SJ. Huntington’s disease: from molecular pathogenesis to clinical treatment. *Lancet Neurol* 2011; 10: 83–98.
- Hayes MT. Parkinson’s disease and parkinsonism. *Am J Med* 2019; 132: 802–807.
- Dexter DT and Jenner P. Parkinson disease: from pathology to molecular disease mechanisms. *Free Radical Biol Med* 2013; 62: 132–144.
- Haaxma CA, Bloem BR, Borm GF, et al. Gender differences in Parkinson’s disease. *J Neurol Neurosurg Psychiatry* 2007; 78: 819–824.
- Wirdefeldt K, Adami HO, Cole P, et al. Epidemiology and etiology of Parkinson’s disease: a review of the evidence. *Eur J Epidemiol* 2011; 26: 1–58.
- Gibb W and Lees A. The relevance of the lewy body to the pathogenesis of idiopathic Parkinson’s disease. *J Neurol Neurosurg Psychiatry* 1988; 51: 745–752.
- Ghosh R and Tabrizi SJ. Clinical features of Huntington’s disease. *Polyglutamine Disord* 2018; 9: 1–28.
- Hart EP, Marinus J, Burgunder JM, et al. Better global and cognitive functioning in choreatic versus hypokinetic-rigid Huntington’s disease. *Mov Disord* 2013; 28: 1142–1145.
- McColgan P and Tabrizi SJ. Huntington’s disease: a clinical review. *Eur J Neurol* 2018; 25: 24–34.
- Zielonka D and Stawinska-Witoszynska B. Gender differences in non-sex linked disorders: insights from Huntington’s disease. *Front Neurol* 2020; 11: 571.
- Talman LS and Hiller AL. Approach to posture and gait in Huntington’s disease. *Front Bioeng Biotechnol* 2021; 9: 668699.
- Landolfi A, Ricciardi C, Donisi L, et al. Machine learning approaches in Parkinson’s disease. *Curr Med Chem* 2021; 28: 6548–6568.
- Fuentes-Ramos M, Sánchez-DelaCruz E, Meza-Ruiz IV, et al. Neurodegenerative diseases categorization by applying the automatic model selection and hyperparameter optimization method. *J Intell Fuzzy Syst* 2022; 42: 4759–4767.
- Mirelman A, Bonato P, Camicioli R, et al. Gait impairments in Parkinson’s disease. *Lancet Neurol* 2019; 18: 697–708.
- Bloem BR, Marinus J, Almeida Q, et al. Measurement instruments to assess posture, gait, and balance in Parkinson’s disease: critique and recommendations. *Mov Disord* 2016; 31: 1342–1355.
- Albani G, Cimolin V, Fasano A, et al. “Masters and servants” in Parkinsonian gait: a three-dimensional analysis of biomechanical changes sensitive to disease progression. *Funct Neurol* 2014; 29: 99.
- Pistacchi M, Gioulis M, Sanson F, et al. Gait analysis and clinical correlations in early Parkinson’s disease. *Funct Neurol* 2017; 32: 28.
- Rovini E, Maremmanni C, Moschetti A, et al. Comparative motor pre-clinical assessment in Parkinson’s disease using supervised machine learning approaches. *Ann Biomed Eng* 2018; 46: 2057–2068.
- Tien I, Glaser SD and Aminoff MJ. Characterization of gait abnormalities in Parkinson’s disease using a wireless inertial sensor system. In: *2010 annual international conference of the IEEE engineering in medicine and biology. IEEE, 2010*. pp.3353–3356.
- Zhan A, Mohan S, Tarolli C, et al. Using smartphones and machine learning to quantify Parkinson disease severity: the mobile Parkinson disease score. *JAMA Neurol* 2018; 75: 876–880.
- Casaca-Carreira J, Temel Y, Van Zelst M, et al. Coexistence of gait disturbances and chorea in experimental Huntington’s disease. *Behav Neurol* 2015; 2015: 970204.
- Keren K, Busse M, Fritz NE, et al. Quantification of daily-living gait quantity and quality using a wrist-worn accelerometer in Huntington’s disease. *Front Neurol* 2021; 12: 719442.
- Waddell EM, Dinesh K, Spear KL, et al. George@: a pilot study of a smartphone application for Huntington’s disease. *J Huntingtons Dis* 2021; 10: 293–301.
- Gaßner H, Jensen D, Marxreiter F, et al. Gait variability as digital biomarker of disease severity in Huntington’s disease. *J Neurol* 2020; 267: 1594–1601.
- Mannini A, Trojaniello D, Cereatti A, et al. A machine learning framework for gait classification using inertial sensors: application to elderly, post-stroke and Huntington’s disease patients. *Sensors* 2016; 16: 134.
- Scheid BH, Aradi S, Pierson RM, et al. Predicting severity of Huntington’s disease with wearable sensors. *Front Digital Health* 2022; 4. DOI: 10.3389/fdgh.2022.874208.
- Zhang S, Poon SK, Vuong K, et al. A deep learning-based approach for gait analysis in Huntington disease. In: *MEDINFO 2019: health and wellbeing e-networks for all*. IOS Press, 2019, pp.477–481.
- Costa L, Gago MF, Yelshyna D, et al. Application of machine learning in postural control kinematics for the diagnosis of Alzheimer’s disease. *Comput Intell Neurosci* 2016; 2016: 3891253.
- Ghoraani B, Boettcher LN, Hssayeni MD, et al. Detection of mild cognitive impairment and Alzheimer’s disease using dual-task gait assessments and machine learning. *Biomed Signal Process Control* 2021; 64: 102249.
- Xia Y, Gao Q, Lu Y, et al. A novel approach for analysis of altered gait variability in amyotrophic lateral sclerosis. *Med Biol Eng Comput* 2016; 54: 1399–1408.

34. Seifollahi M, Mehraban AH, Galvin JE, et al. Alzheimer's disease detection using comprehensive analysis of timed up and go test via Kinect V. 2 camera and machine learning. *IEEE Trans Neural Syst Rehabil Eng* 2022; 30: 1589–1600.
 35. Nam Nguyen QD, Liu AB and Lin CW. Development of a neurodegenerative disease gait classification algorithm using multiscale sample entropy and machine learning classifiers. *Entropy* 2020; 22: 1340.
 36. Jin L, Lv W, Han G, et al. Gait characteristics and clinical relevance of hereditary spinocerebellar ataxia on deep learning. *Artif Intell Med* 2020; 103: 101794.
 37. Summa S, Tartarisco G, Favetta M, et al. Validation of low-cost system for gait assessment in children with ataxia. *Comput Methods Programs Biomed* 2020; 196: 105705.
 38. Vyšata O, Ťupa O, Procházka A, et al. Classification of ataxic gait. *Sensors* 2021; 21: 5576.
 39. Pasquini L, Llibre Guerra J, Prince M, et al. Neurological signs as early determinants of dementia and predictors of mortality among older adults in Latin America: a 10/66 study using the neuroex assessment. *BMC Neurol* 2018; 18: 1–11.
 40. Sánchez-DelaCruz E, Acosta Escalante F, Boll M, et al. Categorización de enfermedades neurodegenerativas a partir de marcadores biológicos de la marcha. *Komputer Sapiens Año VII*: 16, 2015.
 41. Kotthoff L, Thornton C, Hoos HH, et al. Auto-weka 2.0: automatic model selection and hyperparameter optimization in weka. *J Mach Learn Res* 2016; 17: 1–5.
 42. Thornton C, Hutter F, Hoos HH, et al. Auto-weka: combined selection and hyperparameter optimization of classification algorithms. In: *Proceedings of the 19th ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 2013, pp.847–855.
 43. Weka 3-data mining with open source machine learning software in java, <https://www.cs.waikato.ac.nz/ml/weka/> (accessed 28 January 2023).
 44. Breiman L. Random forests. *Mach Learn* 2001; 45: 5–32.
 45. Cleary JG and Trigg LE. K*: An instance-based learner using an entropic distance measure. In: Prieditis A and Russell S (eds) *Machine learning proceedings 1995*. San Francisco (CA): Morgan Kaufmann, 1995, pp.108–114. ISBN 978-1-55860-377-6. <https://doi.org/10.1016/B978-1-55860-377-6.50022-0>. <https://www.sciencedirect.com/science/article/pii/B9781558603776500220>.
 46. Ho TK. The random subspace method for constructing decision forests. *IEEE Trans Pattern Anal Mach Intell* 1998; 20: 832–844.
 47. Panov P and Džeroski S. Combining bagging and random subspaces to create better ensembles. In: *International symposium on intelligent data analysis*. Springer, 2007, pp.118–129.
 48. Murtagh F. Multilayer perceptrons for classification and regression. *Neurocomputing* 1991; 2: 183–197.
 49. Delashmit WH, Manry MT et al. Recent developments in multilayer perceptron neural networks. In: *Proceedings of the seventh annual memphis area engineering and science conference, MAESC*, volume 7, 2005, p.33.
 50. Popescu MC, Balas VE, Perescu-Popescu L. Multilayer perceptron and neural networks, et al. *WSEAS Trans Circuits Syst* 2009; 8: 579–588.
 51. Dorsey ER, Elbaz A, Nichols E, et al. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol* 2018; 17: 939–953.
 52. Santos-Lobato BL, Schumacher-Schuh AF and Mata IF. Lack of full sequencing GBA1 studies for patients with Parkinson's disease in Latin America. *npj Parkinson's Disease* 2022; 8. DOI: 10.1038/s41531-022-00358-z.
 53. Walker RH, Gatto EM, Bustamante ML, et al. Huntington's disease-like disorders in Latin America and the Caribbean. *Parkinsonism Relat Disord* 2018; 53: 10–20.
-