scientific reports

Check for updates

OPEN

Distinguishing amyotrophic lateral sclerosis from radiculopathy using machine learning to analyze nerve conduction data

Armin Ariaei[1,2,3], S. Talebi[4]✉, Bahram Haghi Ashtiani[5], Michael R. Hamblin[6], Mostafa Alipour Langouri[4] & Fatemeh Ramezani[3]✉

Amyotrophic lateral sclerosis (ALS) is a rare, fatal, and irreversible disease that shares some key clinical features with radiculopathy, including muscle atrophy, muscle cramps, and fasciculation. The aim of this study was to find a reliable method to differentiate these two diseases. Machine learning was used to discover new clinical biomarkers for the differential diagnosis of ALS from radiculopathy using nerve conduction study (NCS) data from patients. Data preparation and feature selection were performed by a random forest classifier algorithm, as well as a confusion matrix tool for model selection. After selecting the minimum number of features and the best algorithm, grid search cross-validation was used to optimize the hyperparameters of the chosen algorithm. 77 features were ranked according to their importance. The results of 20 algorithms acting on 8 different groups of features showed that the best performance (accuracy, precision, recall, f-1 score) was obtained using 35 important features and the XGB algorithm, particularly for the recall parameter. Using the XGB algorithm, ALS patients could be identified with accuracy = 0.871, precision = 0.923, recall = 0.850, and f-1 score = 0.857. The XGB algorithm using 35 NCS features could differentiate radiculopathy from ALS in patients with high accuracy.

Keywords Amyotrophic lateral sclerosis (ALS), Radiculopathy, nerve conduction study (NCS), Machine learning

Amyotrophic lateral sclerosis (ALS) is a rare (approximately 5 cases per 100,000 individuals), incurable and fatal disease[1,2]. This progressive disease is characterized by the dysfunction of upper and lower motor neurons[3,4]. The common symptoms of ALS are muscle atrophy and weakness, spasticity, dysphagia, and sometimes psychological disturbance[5]. The term pseudobulbar affect (PSA) has been used to describe the inexplicable sudden bursts of laughter or crying in these patients[6]. Since the main clinical signs of ALS, including muscle atrophy, cramping, and fasciculation[7–9] are common in other diseases such as radiculopathy, a reliable method for differential diagnosis is required[10–13]. The clinical signs of early stages of ALS are similar to those of cervical or lumbar radiculopathy[14,15].

ALS is a clinical diagnosis and requires the exclusion of other disorders that may have similar symptoms, but could have a more favorable prognosis or an effective therapy[16]. For quantifying the disease intensity, the Revised Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R) questionnaire has been developed to rank patients and their symptoms[17,18]. Objective methods used to assess muscle and nerve function in ALS patients, include electromyography (EMG) and nerve conduction study (NCS)[14].

EMG and NCS have been widely used for the differential diagnosis of ALS-like diseases, such as Hirayama disease (HD), spinobulbar muscular atrophy (SBMA), and distal-type cervical spondylitis amyotrophic (CSA)

[1]Men's Health and Reproductive Health Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [2]Student Research Committee, Iran University of Medical Sciences, Tehran, Iran. [3]Physiology Research Center, Iran University of Medical Sciences, Tehran, Iran. [4]Department of Energy Engineering and Physics, Amirkabir university of technology (Tehran Polytechnic), 424 Hafez Avenue, 15875-4413, Tehran, Iran. [5]Department of Neurology, Firoozgar Hospital, School of Medicine, Iran University of Medical Sciences, Tehran, Iran. [6]Laser Research Centre, University of Johannesburg, Doornfontein, Johannesburg, South Africa. ✉email: sa.talebi@aut.ac.ir; framezani2014@gmail.com; ramezani.f@iums.ac.ir

and radiculopathy[19–21]. The primary symptoms of radiculopathy and ALS are similar and there is no accurate method to clinically differentiate between them[22,23].

Radiculopathy is a neurologic disorder characterized by dysfunction of a cervical spinal nerve, nerve root, or both. Radiculopathy can have different symptoms and different names depending on whereabouts in the spine it occurs. When radiculopathy occurs in the lower back, it is known as lumbar radiculopathy, Cervical radiculopathy describes a compressed nerve root in the neck while thoracic radiculopathy refers to a compressed nerve root in the thoracic area of the spine, which affects the upper back. Radiculopathy usually presents as neck and shoulder pain, with a combination of sensory loss, motor function loss, or reflex changes in the region of the damaged nerve root[24–26]. Because there is currently no specific diagnostic biomarker that is sensitive or specific for ALS, the development of a new differential method is interesting.

Currently, machine learning has been applied to analyze the data derived from ALS patients in order to predict disease progression and find new diagnostic biomarkers[27,28]. Random forest (RF) was applied to ALSFRS-R scores, while other machine learning methods have been proposed including, support vector machine (SVM), discriminant analysis, canonical discriminant function, boosting, and artificial neural networks (ANN)[29]. Machine learning was performed on MRI data from ALS patients, using SVM and RF methods due to the limited sample size. Nevertheless, one limitation of these methods is the hidden logic of data processing. In this case, the decision tree is a better interpretation of how the data could be classified based on the chosen variable[30]. Moreover, Bayesian decision trees have been used in some studies because of their better performance in slope prediction[31–33].

In this study, the NCS data from ALS and radiculopathy patients were analyzed to discover novel clinical biomarkers for early-stage ALS detection and differentiation from radiculopathy. These biomarkers could be used when there are insufficient clinical data to make a reasonable diagnosis.

Literature review

Some other studies have been conducted to diagnose ALS based on NCS data. In one study by Hu et al., the NCS data from ALS patients with lower limb dysfunction were investigated. In that study, a "split leg index" was defined as the dorsiflexion CMAP value divided by the plantar flection CMAP value. Correspondingly, the study reported 0.41 (AUC = 0.743, 95% CI 0.682–0.805) as the cutoff value of split leg index to differentiate ALS patients from normal cases, and a value of 0.52 (AUC = 0.749, 95% CI 0.602–0.896) to differentiate ALS from lumbar spondylosis[34].

In another study conducted by Jing Fan et al. in 2023, the differential diagnosis of ALS from radiculopathy was described based on the severity of fasciculation assessed by muscle ultrasound and the Medical Research Council (MRC) scale. Values were compared in the two diseases, with the results showing that combined analysis of fasciculation intensity and MRC score of limb muscles could be useful for the differential diagnosis of ALS[35].

A study by Tannemaat et al. in 2022, created an automated time series classification algorithm for distinguishing normal, neuropathic, and EMG traces. A machine learning pipeline was applied based on 5-s EMG fragments of each muscle. The results showed that an automated time series classification algorithm could distinguish between EMGs from healthy individuals and those from patients with ALS with a high diagnostic yield[36].

In another study conducted by Kalita et al. in 2017 on ALS patients, they investigated the ability of the upper extremity NCS to differentiate ALS from Hirayama disease. They employed a split hand index (SHI) and the ulnar to median ratio. SHI was defined as the multiplication product of the abductor pollicis brevis CMAP with the first dorsal interosseous CMAP and then divided by the abductor digiti minimi CMAP. To detect ALS disease they chose an ulnar to median ratio > 1.2, with a sensitivity of 60% and specificity of 90.2%. When they chose an SHI ≤ 5.2 there was a sensitivity of 82% and a specificity of 88.8%[37].

Similarly, in a study by Menon et al., SHI was calculated in 44 ALS patients, and by applying a cutoff point of 5.2, the sensitivity and specificity were calculated to be 74% and 80%, respectively[38]. In a recent article by Souayah et al., the mechanism underlying the decreased conduction velocity was explained by destruction of large axons. Accordingly, they suggested that the median nerve would be the most vulnerable to loss of conduction velocity due to demyelination[39]. In this study the features obtained from the median nerve were among the most important to differentiate ALS from radiculopathy.

Furthermore, in a study by García et al., an estimation of the motor unit number was suggested as a variable to differentiate ALS from other diseases that mimic ALS symptoms. Accordingly, by setting the cut-off point of the motor unit number at 0.548, there was a sensitivity of 82% and a specificity of 75% for the detection of ALS disease[40]. In another study including 19 patients with ALS, the split hand index (SHI) was calculated and compared to the normal group. The study reported a sensitivity, specificity, and accuracy of 72%, 94%, and 83%, respectively, by setting the cut-off point at 7.2[41].

So far, most studies have attempted to diagnose ALS based on the relationship between various features obtained from NCS, but in our study, we carried out feature selection with Mean decrease impurity to choose the best model for ALS diagnosis.

Methods

IRAN University of Medical Sciences ethics approval center approved the study, including all relevant details; and all experiments were performed in accordance with this committee's guidelines.

Subjects

This was a single-center retrospective study on patients referred to the Firoozgar Hospital (located in Tehran, Iran) from January 2018 to March 2023 who received NCS tests (a total of 6636 individuals) (Sample data in Supplementary file2).

In the ALS patients, 34.6% were female and 65.4% were male. In the radiculopathy patients, 48.4% were female and 51.6% were male. All patients were Iranian nationality. The age of the patients ranged from 35 to 70 years. This age range was chosen because people under 35 years of age mostly suffer from another form of this disease, known as Juvenile Amyotrophic Lateral Sclerosis (JALS), and its hereditary and familiar characteristics, and pathological mechanisms are different from the common type of ALS. On the other hand, people over 70 years of age in the general population often have underlying diseases such as diabetes that can cause diabetic neuropathy, or have symptoms of carpal tunnel syndrome due to their older age, which could be similar to the split hand seen in ALS. Patients with only upper or lower extremity data, patients with missing data, and patients with carpal tunnel syndrome (CTS) in addition to ALS were excluded. The criteria used to identify CTS was a significant reduction in the velocity (below 47 m/s) of the sensory division of the median nerve together with a normal sensory conduction velocity in the ulnar nerve[41]. Finally, patients who were diagnosed with radiculopathy but did not have ALS were selected as a group to differentiate the symptoms of ALS. Radiculopathy cases included cervical polyradiculopathy, lumbosacral polyradiculopathy, CTS and cervical polyradiculopathy. ALS patients were diagnosed based on the El Escorial ALS criteria with subgroups of definitive ($n = 114$) and probable ($n = 35$). Data from the EMG testing of muscles in the bulbar, cervical, thoracic, and lumbosacral regions indicating neurogenic damage (giant/polyphasic motor unit potentials, spontaneous muscle fiber activity, and reduced recruitment) were used to differentiate ALS from other clinically similar conditions[40]. After applying these inclusion criteria, 149 ALS, 378 radiculopathy, and 130 axonal sensory-motor polyneuropathy patients remained in the study.

NCS procedure

NCS was conducted on normal skin at a temperature of above 32° on four motor and sensory nerves, two in the upper extremity (median and ulnar) and two in the lower extremity (peroneal and tibial, sural and superficial), using UltraPro® S100 EMG/NCS/EP Neurodiagnostic System (Natus, Wisconsin, USA) (Supplementary file1, Figure 1S). There was a 2–3 cm distance between the reference and recording electrodes. NCS assessment was conducted by recording electrodes over an approximately 6 cm distance for the upper extremity and 12–14 cm for the lower extremity[42]. However, the location of the stimulator and recording electrodes were adjusted to obtain the best amplitude of the CMAP. The motor conduction data was obtained from the median nerve innervating *abductor pollicis brevis* (APB) muscle, the ulnar nerve innervating *abductor digit minimi* (ADM) muscle, the peroneal nerve innervating *extensor digitorum brevis* (EDB) muscle, and the tibial nerve innervating *abductor hallucis* (AH) muscle. The antidromic technique was used to obtain sensory NCS. The stimulation location was the same as motor NCS, but the recording electrodes were located in different places. Accordingly, in the upper extremity, the recording electrode was placed on the interphalangeal joint of the second and fifth digits for median and ulnar nerves, respectively. Moreover, in the lower extremity, the recording electrode was placed inferior to the lateral malleolus and 2 finger-widths anterior to the lateral malleolus for sural and superficial nerves[43]. The distal latency, amplitude, and velocity of the SNAP and CMAP were calculated and analyzed as described in previous studies[44,45] (Fig. 1).

Study design

Firstly, the data was labeled. Individuals with ALS disease were labeled as positive (+), and other cases were labeled as negative (−). Considering that the labels are known, supervised algorithms should be used to separate the labels. The chosen algorithm must find connections between the features in order to perform a classification based on the labels. Considering that the number of features is large, the problem is clearly non-linear, and finding boundaries between groups is not easy. Therefore, feature selection should be done to reduce the complexity of the problem and also avoid overfitting. Moreover, a suitable algorithm for dealing with this problem should be selected (Supplementary file 1, Figure 2S).

Mean decrease impurity feature importance

- When machine-learning algorithms are applied to high-dimensional data, they face a critical issue known as the "curse of dimensionality". This phenomenon occurs because data become sparser in high-dimensional space, adversely affecting algorithms designed for low-dimensional space. The curse of dimensionality leads to several significant challenges[46]Sparsity of Data: As the dimensionality increases, the volume of the space increases exponentially, causing data points to become sparse. This sparsity makes it difficult for algorithms to find meaningful patterns and relationships within the data.
- Overfitting
- Increased Memory Requirement and Computational Costs

Effective feature selection and dimensionality reduction techniques are often employed to mitigate the curse of dimensionality, enhance model performance, and reduce computational complexity. For this reason, the Mean Decrease Impurity (MDI) feature importance[47] method has been used.

The MDI method was chosen due to its suitability for high-dimensional datasets and interpretability in medical data applications[48–50]. This method was selected for several reasons. Firstly, Random Forests are inherently well-suited to handle high-dimensional data and complex interactions among predictors without strong assumptions about data distribution. Secondly, the MDI metric provides an intuitive measure of feature importance by quantifying how much each predictor reduces impurity across all trees in the ensemble. This property simplifies interpretability and helps to decide which features have the greatest impact on classification performance. Random Forest feature selection was chosen due to its balance of simplicity, interpretability, and robustness across a variety of data scenarios.

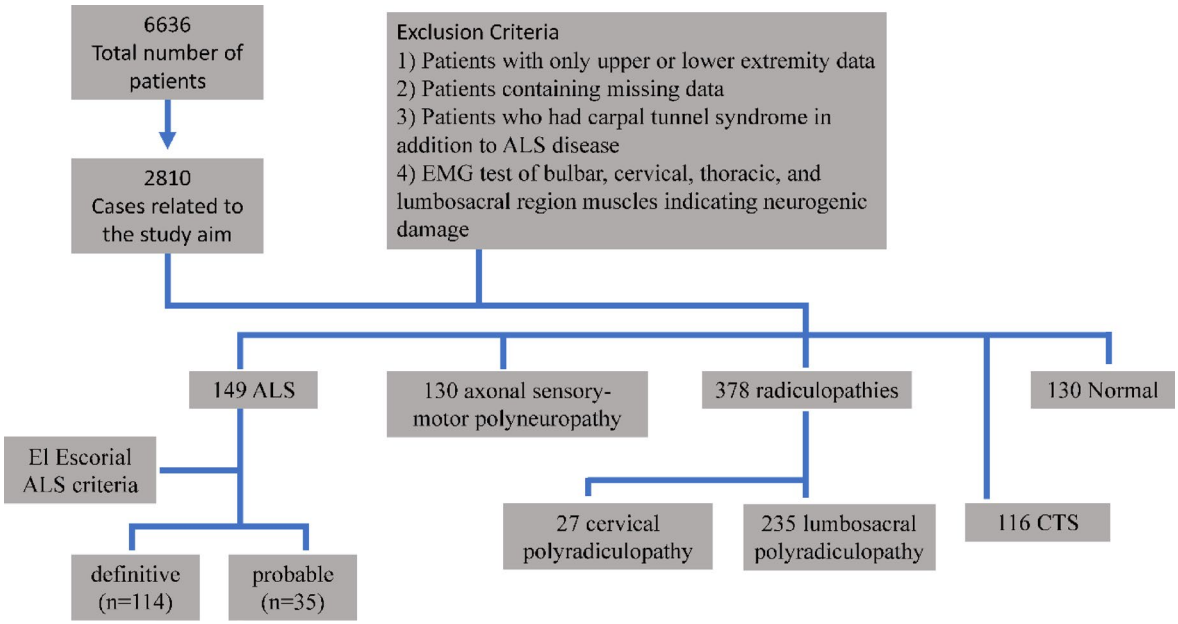

Fig. 1. The data-screening workflow in this study to obtain the appropriate samples. During 6 years a total of 6636 individuals received NCS testing in Firoozgar hospital. After applying the inclusion criteria, 149 ALS, 378 radiculopathy, and 130 axonal sensory-motor polyneuropathy patients remained in the study. ALS patients diagnosed with the El Escorial ALS criteria with subgroups of definitive ($n = 114$) and probable ($n = 35$). Radiculopathy patients were subdivided into lumbosacral polyradiculopathy (235) and cervical polyradiculopathy (27).

MDI operates effectively with tree-based algorithms, particularly ensemble methods like XGBoost, which can robustly handle large datasets with numerous features. The use of MDI has been validated in similar studies on high-dimensional medical datasets, where the trade-off between computational efficiency, interpretability, and robustness can be critical. Therefore, MDI was a logical choice in this study to ensure a balance of performance, accuracy, and clarity in feature selection.

When constructing decision trees and random forests, an essential task is determining the quality of a split at any given node. This quality is measured by the decrease in impurity, denoted as Δi(s,t). The decrease in impurity Δi(s,t) for a node t split using a split sss is mathematically expressed as:

$$\Delta i\left(s,t\right)=i\left(t\right)-p_{L}i\left(t_{L}\right)-p_{R}i\left(t_{R}\right)$$

where:

- $i\left(t\right)$ represents the impurity of the node t before the split.
- $i\left(t_{L}\right)$ is the impurity of the left child node t_{L} after the split.
- $i\left(t_{R}\right)$ is the impurity of the right child node t_{R} after the split.
- p_{L} is the proportion of samples that fall into the left child node t_{L} after the split, it was calculated as $p_{L}=\frac{N_{t_{L}}}{N_{t}}$, where $N_{t_{L}}$ is the number of samples in t_{L}.r of samples in the parent node t.
- p_{R} is the proportion of samples that fall into the right chil. node t_{R} after the split, it was calculated as $p_{R}=\frac{N_{t_{R}}}{N_{t}}$, where $N_{t_{R}}$ is the number of samples in t_{R} and N_{t} is the number of samples in the parent node t.

Impurity measures are designed to quantify the level of disorder or impurity within a node. Common impurity measures include.

- Gini Index: Often used in classification tasks, it measures the probability of misclassifying a randomly chosen element if labeled according to the distribution of labels in the node.
- Shannon Entropy: Another measure used in classification to quantify the amount of uncertainty or randomness in the labels.
- Variance: Typically used in regression tasks, it measures the dispersion of the target variable values within the node.

The impurity value measures how mixed the labels are within a node. The goal is to find splits that produce the most homogeneous child nodes.

The importance of a variable X_m for predicting Y is evaluated by adding up the weighted impurity decreases $p(t)\,\Delta i(s_t, t)$ for all nodes t where X_m is used, averaged over all N_T trees in the forest. The importance of variable X_m can be calculated as follows:

$$\mathrm{Imp}\left(\mathrm{X_m}\right) = \frac{1}{\mathrm{N_T}} \sum\nolimits_{\mathrm{T}} \sum\nolimits_{\mathrm{t \in T: v(s_t) = X_m}} \mathrm{p\,(t)\,\Delta i\,(s_t, t)}$$

where:

- N_T is the number of trees in the forest.
- t is node in tree T.
- $v(s_t) = X_m$ indicates that the variable X_m is used in t split s_t.
- $p(t)$ is the proportion $\frac{N_t}{N}$ of samples reaching node t.
- $\Delta i(s_t, t)$ is the impurity decrease resulting from the spl at node t.

Model selection

The dataset contained 787 numerical records, which is relely small for training Deep Learning models effectively. Deep Learning techniques generally require larger datasets to prevent overfitting and ensure robust generalization, particularly for complex models like neural networks. The use of machine learning allowed us to leverage the available data more effectively.

In this study, the dataset showed a class imbalance between A and non-ALS cases, with ALS cases being relatively rare compared to non-ALS cases. This imbalance could create challenges for machine learning models, because models trained on imbalanced data may become biased toward the majority class (non-ALS), leading to lower sensitivity in identifying ALS patients (true positives).

To address this issue, we employed two strategies.

1. Stratified Cross-Validation: By using stratified k-fold cross-validation, we ensured that each fold maintained the same class proportions as the original dataset. This approach helped the model to learn from a more balanced representation of classes across all folds, reducing the risk of bias toward the majority class.
2. Evaluation Metrics: In addition to overall accuracy, we focused on additional evaluation metrics such as precision, recall, and F1-score, especially for the ALS class (positive class). These metrics provide a more nuanced view of model performance, particularly in cases where accuracy alone could be misleading due to the class imbalance.

These techniques collectively helped to mitigate the impact of data imbalance on model performance, ensuring that our models were capable of accurately identifying ALS cases despite the disparity in class representation.

After finding the most important features, in order to determine the lowest number of features with the best accuracy for the machine learning algorithm, training was first carried out on a limited number of features with the highest importance, and at each stage, additional features were added in order of importance and the accuracy of the algorithm was checked again. The accuracy should not change from one step to the next. The final step is the stage where the optimum number of features is used. Therefore, the data processing engineer must choose the most appropriate algorithm for training a specific database and should not be satisfied with a limited number of algorithms (Table 1).

To evaluate the accuracy of these algorithms and identify the best-performing model, we used the confusion matrix tool in Python. Initially, we performed a hold-out test by using 10% of the dataset as a test set to get preliminary performance metrics. However, to improve confidence in the results, we further applied stratified cross-validation. Stratified cross-validation was chosen to ensure that each fold maintains the same class proportion as the original dataset, which is critical in datasets with imbalanced classes, like ALS vs. non-ALS cases.

Specifically, we implemented a stratified k-fold cross-validation with k = 10 to better assess model robustness and to reduce the variance in performance metrics. Each model's performance was averaged across the cross-validation folds, providing a more reliable estimate of its accuracy and stability. The confusion matrix shown in Table 2 reflects the output of the cross-validation process, where TP represents the number of ALS patients correctly identified across all folds, FP represents non-ALS cases incorrectly predicted as ALS, FN represents ALS patients incorrectly classified as non-ALS, and TN represents non-ALS cases correctly classified as non-ALS.

In general, to check the performance of a supervised machine learning algorithm, four parameters can be calculated by the confusion matrix tool.

$$1.\ Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Accuracy was calculated to evaluate whether the algorithm could correctly differentiate the two groups (ALS cases and non-ALS cases).

$$2.\ Precision = \frac{TP}{TP + FP}$$

Precision is the ratio of the number of patients that the algorithm correctly predicted, to the total number of patients included. If the number of false positives that the machinpredicts is low, the numerical value of

Number	Supervised machine learning algorithm classifiers
1	XGB
2	MLP
3	LogisticRegression
4	SGD
5	PassiveAggressive
6	DecisionTree
7	ExtraTree
8	LinearSVC
9	SVC
10	GaussianNB
11	AdaBoost
12	Bagging
13	RandomForest
14	GaussianProcess
15	GradientBoosting
16	LinearDiscriminantAnalysis
17	QuadraticDiscriminantAnalysis
18	Ridge
19	KNeighbors
20	ExtraTrees

Table 1. The machine learning algorithms utilized in this study.

	Prediction	
Actual	Positive	Negative
Positive	True positive (TP)	False negative (FN)
Negative	False positive (FP)	True negative (TN)

Table 2. Outcomes of the confusion matrix. Positive, diseases; Negative, No diseases. TP represents the number of ALS patients correctly identified across all folds, FP represents non-ALS cases incorrectly predicted as ALS, FN represents ALS patients incorrectly classified as non-ALS, TN represents non-ALS cases correctly classified as non-ALS.

Precision becomes higher and closer to one. Therefore, the closer the precision to one, the fewer non-ALS cases are mistakenly diagnosed as ALS.

$$3.\ Recall\ (Sensitivity) = \frac{TP}{TP + FN}$$

If the algorithm mistakenly identified many FN cases or the number of people who suffered from the actual disease was small, then the recall parameter would be close to one. In other words, recall is the same as sensitivity, and it is the most clinically important parameter in a machine algorithm applied in medicine.

$$4.\ F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

The F1-score is a combination of Precision and Recall.

This study had three main aims. A) To find the minimum number of features for the algorithm to have sufficient accuracy. B) To identify the best machine learning algorithm to achieve accuracy, precision, recall, and f-1 score parameters. C) To hyper-optimize the algorithm parameters.

To achieve the goals, the following items were included in the Python code. (1) To select the optimum number of features, the five most important features were isolated to undergo all 20 algorithms to construct the confusion matrix parameters. By adding further groups of five features to the previous features, the number of features was steadily increased and confusion matrix parameters were obtained for all the algorithms at each step. (2) Due to the fact that the training and test data were randomly selected in different runs, these data may be different from each other, and it would not be possible to make an accurate judgment about their performance. Therefore, the entire data set was divided into ten groups and the cross-validation method available in the scikit-learn subpackage was employed. (3) To prevent information leakage between training and test data, scaling operations,

training, and tests were performed in a pipeline. (4) After selecting the minimum number of features and the best algorithm, grid search cross-validation was used to optimize the hyperparameters of the selected algorithm.

XGBoost

XGBoost or Extreme Gradient Boosting is a decision tree ensemble based on gradient boosting introduced by Chen and Guestrin[50]. It builds multiple decision trees sequentially, where each new tree attempts to correct the errors made by the previous tree. This process, known as gradient boosting, helps the model to learn complex relationships within the data by minimizing a defined loss function at each step.

The key strengths of XGBoost include its efficiency, scalability, and regularization capability, which reduce overfitting and improve generalizability. XGBoost also offers additional features, such as tree pruning and optimal split finding, making it particularly effective on structured, high-dimensional datasets.

In this study, we tested various machine learning models and found that XGBoost could outperform other models in its predictive accuracy on our dataset. Its ability to handle missing values, along with efficient memory usage and parallel processing, made it a suitable choice given the nature of our data and goals.

Code Availability

The code and data used to produce the results of this article is available at the following address. https://github.com/framezani2014-cyber/ALS--COD.git.

Results

The primary data included 787 samples and 78 features (Supplementary file1, Table 1S). Data from the participants were used to test the accuracy of the algorithm. Moreover, "+" and "–" signs were used for coding labels (Table 3).

Feature selection

The features were ordered based on their importance. The result showed the important features in differentiating ALS cases were: RUA = 0.048, RMA = 0.048, RMV = 0.037, LMV = 0.035, LMA = 0.034, RMNP = 0.032, LMPP = 0.025, RSP P = 0.023, RMPP = 0.02, RSPL = 0.019 (Table 4). As can be seen in Table 4, gender is considered as one of the features and has the lowest rank in terms of importance in feature selection.

Training was performed first with a limited number of the most important features. At each stage, other features were added in order of importance, and the accuracy of the algorithm was checked again (Fig. 2).

The 35 most important features were plotted in relation to each other (Fig. 3). The blue plots along the diagonal show the distribution of features among different patients, and the figures in places other than the main diagonal show the correlation between different features. As shown in Fig. 3, the plots in the main diagonal show a Maxwellian distribution, which indicates an appropriate selection of features and their correct measurement. The correlation between features in the non-main diagonal shows that there is only a weak correlation between them and the features are independent of each other.

The performance parameters (accuracy, precision, recall, f-1 score) obtained from the different algorithms were plotted for different numbers of features (5, 15, 25, 35, 45, 44, 65, 77). Accordingly, the optimum number of features was found to be 35. It can be seen that the XGB algorithm showed the best performance, especially for the Recall parameter (Fig. 4). The raw results are available in Supplementary file 1, Table 2S.

Model optimization

The best XGB hyperparameters were obtained using the grid search validation subpackage as follows: Max_depth = 4, learning rate = 0.07, n_estimators = 800, subsample = 0.8.

The performance of XGB was plotted for different numbers of features. It can be seen that selecting 35 features resulted in the best performance with values of accuracy = 0.871, precision = 0.923, recall = 0.850, and f-1 score = 0.857 (Fig. 5). The model achieved an average accuracy of 0.869 (± 0.015), precision of 0.922 (± 0.012), recall of 0.849 (± 0.013), and F1-score of 0.857 (± 0.011) across the 10 folds. These results demonstrate strong and consistent performance, suggesting that the XGB model is robust and can generalize effectively to deal with new data. To mitigate the risk of overfitting, especially considering the dataset's size and complexity, we implemented several robust validation strategies. Firstly, we employed tenfold cross-validation, ensuring that each fold provided a comprehensive evaluation of the model's performance across different data partitions. Additionally, the dataset was divided into distinct training, validation, and testing sets. For each fold, we carefully monitored the accuracy across these three subsets and confirmed that their performance metrics were consistent. This multi-faceted approach indicated that the model was generalizing well rather than simply memorizing the training data, thereby reducing the risk of overfitting.

Normal	ALS Disease	ALS Suspicious
130	149	508
	Positive (+)	Negative (–)

Table 3. Data of 787 cases were categorized into three groups.

Number	Features	Importance
1	RUA	0.048
2	RMA	0.048
3	RMV	0.037
4	LMV	0.035
5	LMA	0.034
6	RMNP	0.032
7	LMPP	0.025
8	RSPP	0.023
9	RMPP	0.02
10	RSPL	0.019
11	RUV	0.018
12	LUA	0.018
13	LUPP	0.018
14	LPA	0.017
15	LMNP	0.017
16	RTL	0.016
17	RUPP	0.016
18	LSFOL	0.016
19	RPA	0.015
20	RSFPL	0.015
21	RSFOL	0.015
22	RSNP	0.015
23	LUL	0.015
24	LSFPL	0.015
25	RUL	0.013
26	LMV	0.013
27	RMD	0.012
28	RSOL	0.012
29	LSFNP	0.012
30	RUNP	0.011
31	LMOL	0.011
32	RPLD	0.01
33	RUD	0.01
34	LSFV	0.01
35	RML	0.01
36	RSV	0.01
37	LTA	0.01
38	LTL	0.01
39	LPLD	0.01
40	LPL	0.01
41	LMD	0.01
42	LSFV	0.01
43	LSOL	0.01
44	LSV	0.01
45	LUV	0.01
46	LMPL	0.01
47	RTA	0.009
48	RPD	0.009
49	RSFV	0.009
50	RSFNP	0.009
51	RMPL	0.009
52	RMOL	0.009
53	RTD	0.009
54	LULD	0.009
55	LSFPP	0.009
56	LSPP	0.009
Continued		

Number	Features	Importance
57	LUNP	0.009
58	RPL	0.008
59	LPV	0.008
60	LPD	0.008
61	LMLD	0.008
62	LSPL	0.008
63	RUVD	0.007
64	RMLD	0.007
65	RSFPP	0.007
66	LUD	0.007
67	LML	0.007
68	LSNP	0.007
69	LUOL	0.007
70	RPV	0.006
71	RULD	0.006
72	RUPL	0.005
73	RUOL	0.005
74	LTD	0.005
75	RTD	0.004
76	LUPL	0.003
77	Gender	0.001

Table 4. The importance of features for the detection of ALS patients.

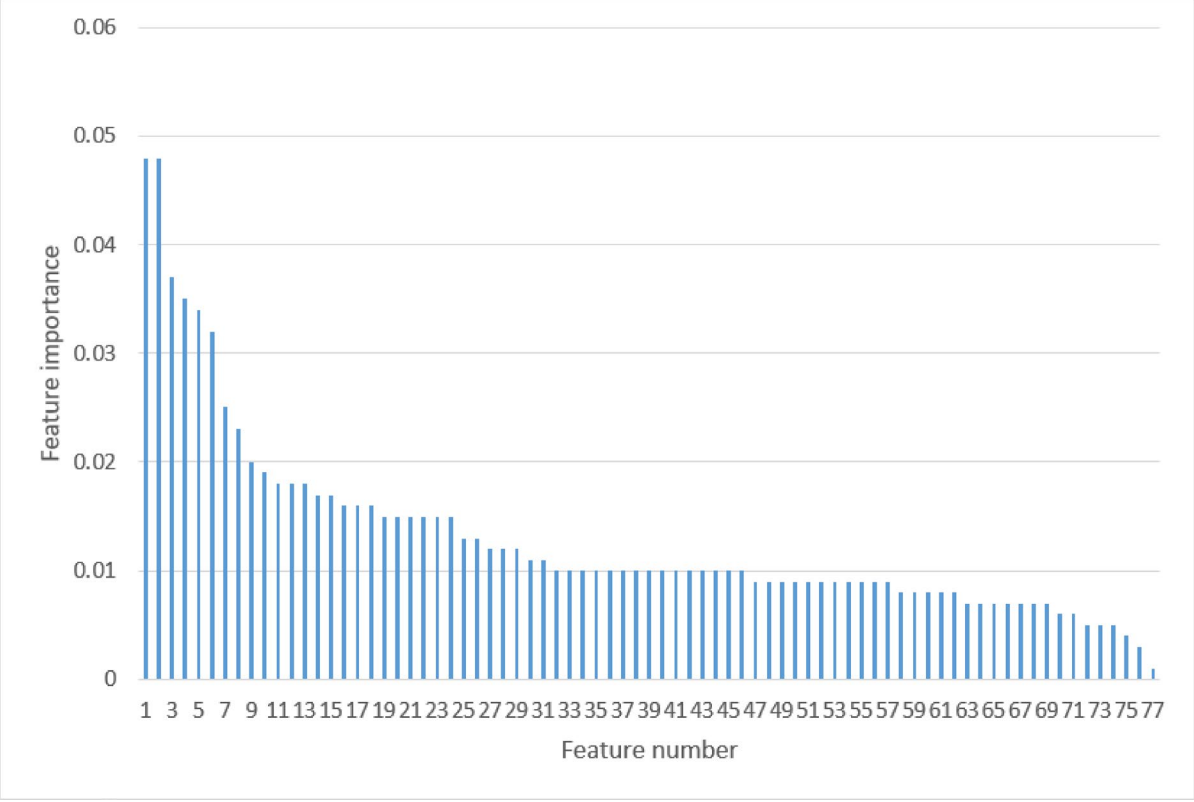

Fig. 2. The distribution of feature importance was plotted based on data from Table 4.

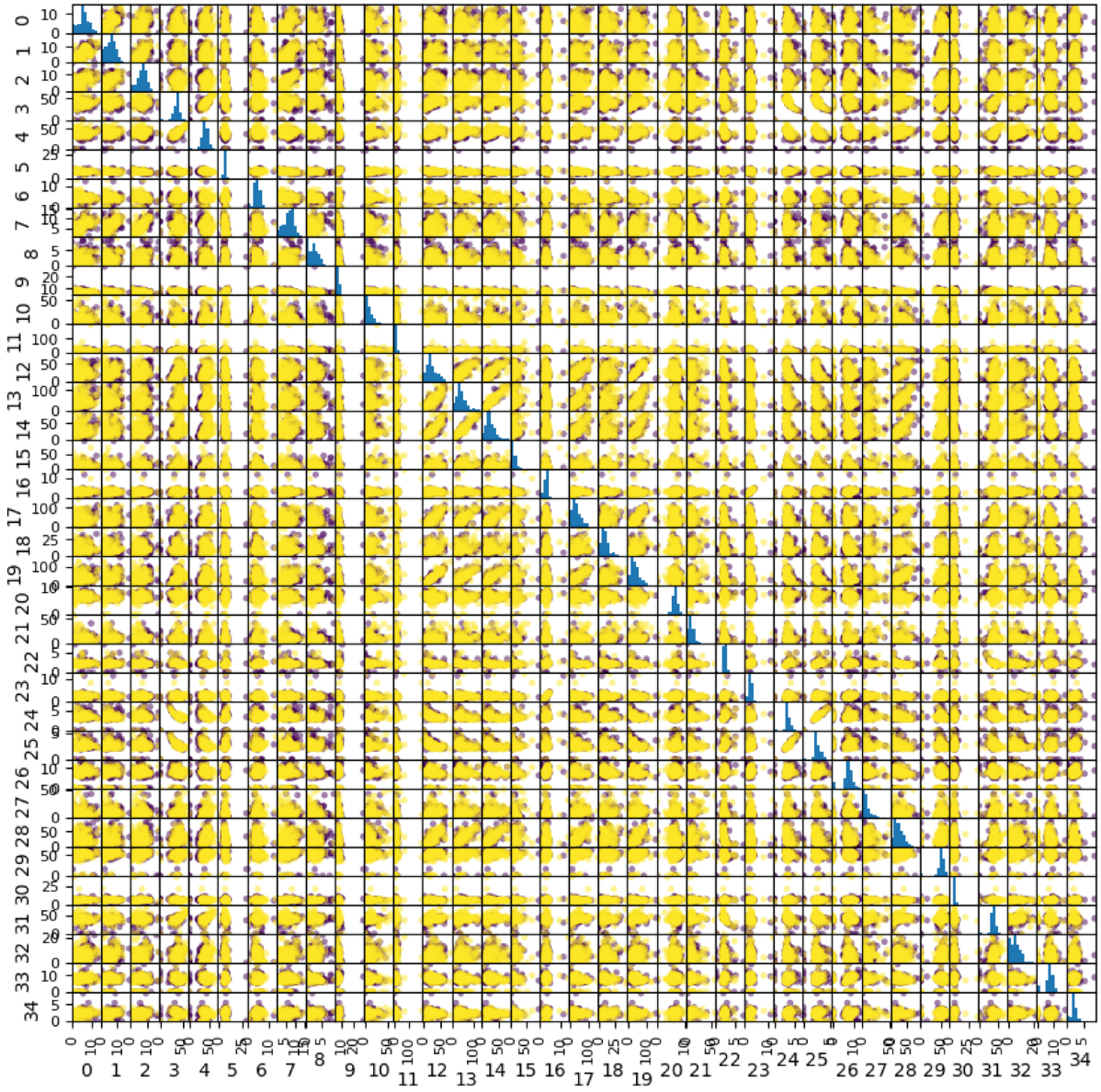

Fig. 3. Showing the distribution of 35 features with the most importance. The labels are marked with yellow and brown dots.

Differential diagnosis

The data of polyradiculopathy patients were classified based on the sensory and motor neuron impairments, into axonal sensory-motor polyneuropathy, and carpal tunnel syndrome (CTS), both classified as "sensory polyradiculopathy", including 246 patients. Lumbosacral polyradiculopathy and cervical polyradiculopathy were classified as "motor polyradiculopathy" including 262 patients. The results of the analysis for the differential diagnosis of each class of radiculopathy from ALS using the XGB algorithm are shown in Table 5. This type of grouping eliminates the possibility of overestimating the discriminating power of the model and shows that the model has the ability to differentiate ALS from all types of radiculopathy with almost the same power in both groups. The ROC curve is shown in Fig. 6.

Discussion

In this study, we employed a supervised learning approach using a random forest classifier algorithm to select the most useful features contained in the NCS data. To choose a simple final model and avoid over or under-prediction, it would be better to use the lowest number of features that did not affect the final accuracy of the results. The results of 20 algorithms operating on groups with 8 different numbers of features (5–77) showed that the best performance (accuracy, precision, recall, and f-1 score) was obtained using the XGB algorithm with the 35 most important features, giving accuracy = 0.871, precision = 0.923, recall = 0.850, and f-1 score = 0.857.

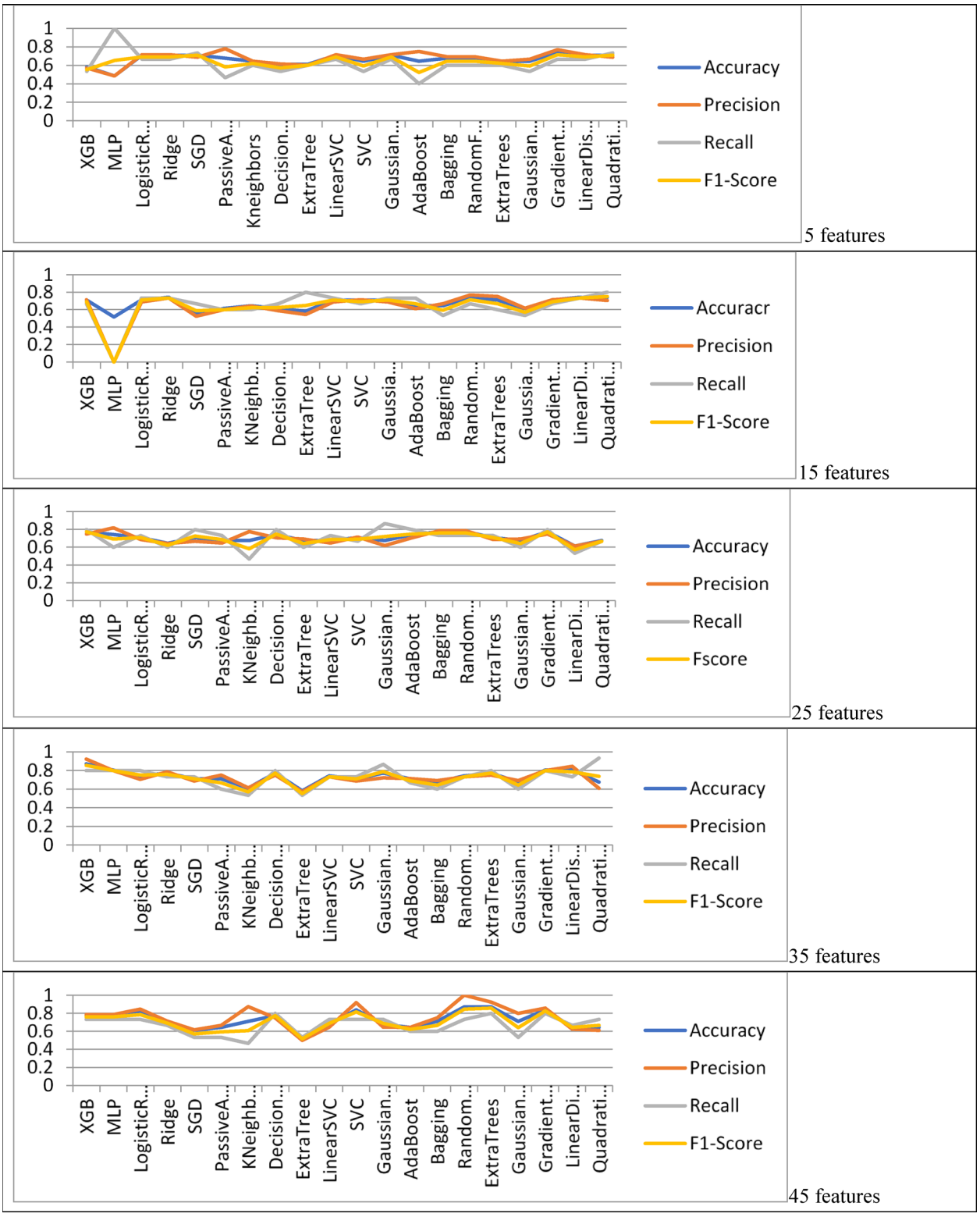

Fig. 4. Accuracy, precision, recall, and f-1 score parameters for different numbers of features (5, 15, 25, 35, 45, 44, 65, 77) obtained from all 20 algorithms.

To address the inherent class imbalance in our dataset, we integrated the Synthetic Minority Over-sampling Technique (SMOTE) into our data preprocessing pipeline. SMOTE generates synthetic examples for the minority class by interpolating between existing samples, rather than merely duplicating them. This approach effectively balances the class distribution, mitigating the model's tendency to favor the majority class. In our study, SMOTE was applied solely to the training set to preserve the natural distribution in the validation and test sets, ensuring

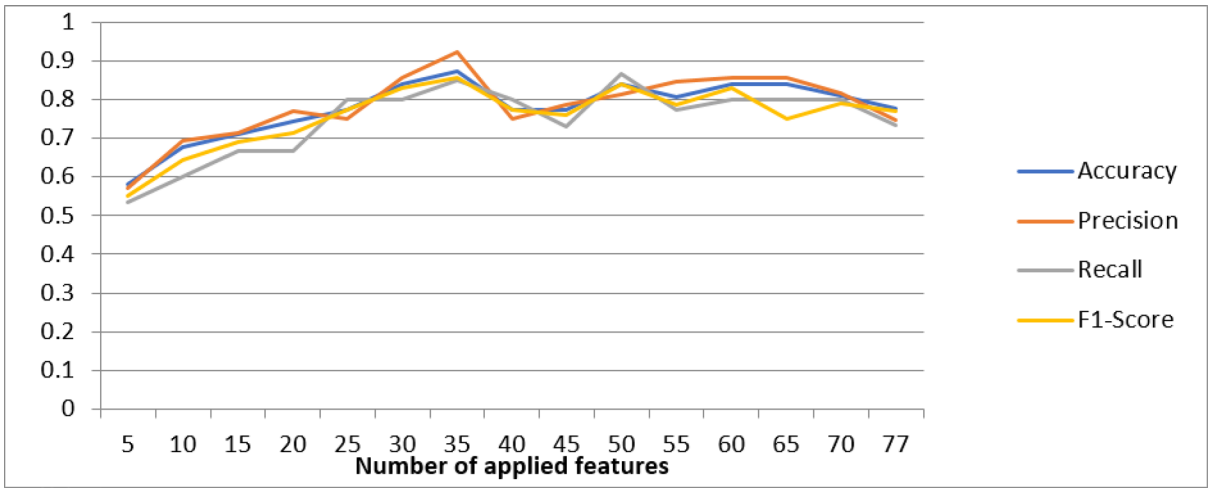

Fig. 5. The performance of the XGB algorithm for different numbers of features.

XGB algorithm	Differential diagnosis of ALS from both kinds of radiculopathy	Differential diagnosis of ALS from motor neuron radiculopathy	Differential diagnosis of ALS from sensory neuron radiculopathy
Accuracy	0.828	0.81	0.84
Precision	0.871	0.833	0.873
Recall	0.923	0.873	0.91
F1-Score	0.8	0.773	0.791

Table5. The performance of the XGB algorithm for differential diagnosis of ALS from motor neuron radiculopathy, sensory impairment radiculopathy and both conditions.

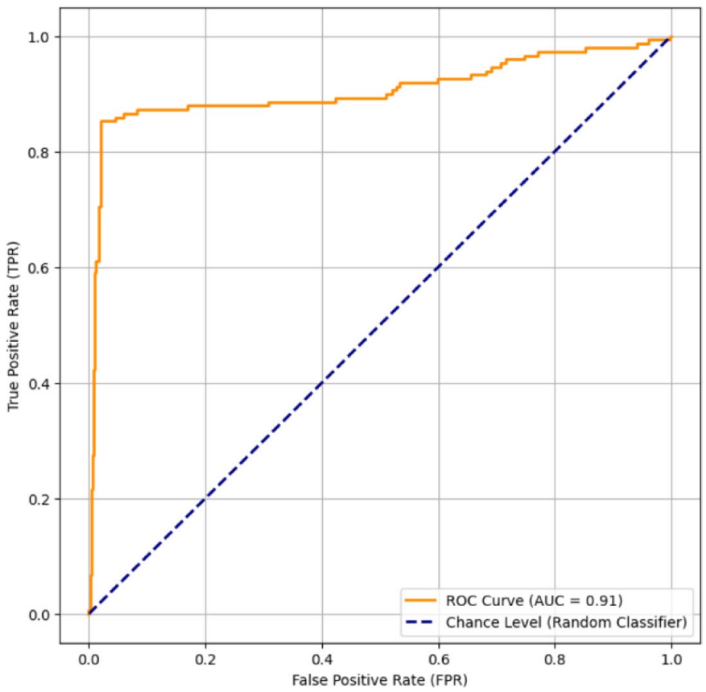

Fig. 6. ROC curve of XGBoost model.

an unbiased evaluation. This technique not only enhanced the model's sensitivity to the minority class but also contributed to more robust and balanced performance metrics, particularly in terms of recall and precision.

The average of each feature in the group of patients with ALS and radiculopathy was also statistically compared, and the values of RUA, RA, RMV, LMA,LUA and MLV showed moderate statistical significant differences (Supplementary file 1, Figure 3S). These features are the ones that showed the greatest effect size after feature selection. However, it should be noted that simple statistical testing compares the average of features in two groups, but it cannot predict with high accuracy individual cases. Therefore, one advantage of machine learning compared to simple statistical analysis, is that it could make predictions about individual patients with more than 90% accuracy in our study.

XGBoost is an algorithm that has recently become more important in applied machine learning and Kaggle competitions for structured or tabular data. XGBoost is also an implementation of gradient-boosted decision trees designed for speed and performance[51].

The reason why XGBoost was chosen as the preferred classification model when dealing with problems encountered in real-world classification is that XGBoost is approximately 10 times faster than existing methods running on the same platform. Therefore the issue of time consumption is minimized when network data preprocessing is carried out. XGBoost also has the advantage of parallel processing, and it uses all the cores of the device it is running on[50,52,53].

The flexibility provided by XGBoost is high and it is not dependent on a specific platform; hence IDS using XGBoost is platform-independent which is a big advantage. XGBoost can be used by several programming languages, such as Java, Python, R, and C + + allowing the use of environments such as parallel computing. XGBoost has the ability to transform a weak learner into a strong learner (boosting). Through its optimization step for each new tree that is connected, the classification model will produce fewer false alarms, allowing easy data labeling, and accurate data classification[51].

Regularization is also an important aspect of the XGBoost algorithm, as it helps to avoid problems of overfitting the data, whether they are tree or linear models. XGBoost is also well-equipped to detect and deal with missing values. All these important features add up and enable XGBoost to outperform many existing models. It can be said that XGBoost is well-suited to deal with most problems that a real-world network could face[54,55].

In our results, 35 important features were identified, some of which had been considered important features in studies by other researchers. The features that have shown the most importance are RUA, RMA, RMV, LMV, and LMA which are related to the amplitude and velocity of ulnar and median nerves. The median and ulnar nerves run through the forearm and wrist, and they help to coordinate the motion of our forearm and hands. The ulnar nerve derives mainly from the C8 and T1 roots and is a minor component of the C7 root[56]. The median nerve sends touch, pain, and temperature signals from the lower arm and hand to the brain. Almost all ulnar nerve fibers pass through the lower trunk and medial cord of the brachial plexus, and then the nerve descends through the upper arm and passes from anterior to posterior through the medial intermuscular septum[57]. Ulnar[57–59] and median[58,60] nerve amplitude has also been reported in other studies to be an important factor in the diagnosis of ALS and the monitoring of disease progression[61]. We observed in this study that the amplitude of the studied nerve was a much more important determinant than the velocity. Nerve conduction velocity mainly reflects changes in the myelin sheath, while the amplitude of the action potential serves as an marker of axonal changes and the condition of the nerve fiber. A reduction in the action potential amplitude indicates axonal damage and provides an estimate of the number of nerve fibers activated by electrical stimulation[62].

Antunes et al. performed a study for the differentiation of ALS from control patients, based on surface electromyography (sEMG). Their validation study with sEMG contained 17 ALS patients and 24 control subjects, leading to the proposal of a set of novel features for sEMG based on signal morphology[63]. Other studies have reported that a reduction in conduction velocity could occur in ALS patients[37], but so far, no study has examined the effect of NCS features individually and their correlated effect in ALS detection.

In CTS and sensory-motor axonal polyneuropathy, sensory symptoms are mostly reported, while in lumbosacral polyradiculopathy and cervical polyradiculopathy, motor symptoms are mostly reported. Because in ALS movement-related symptoms are dominant, it could be that considering all these diseases together might give a wrong estimae of the model's detection power. However, in our model, although it showed the power to distinguish ALS from sensory radiculopathy a little more accurately than from motor radiculopathy, this difference was not significant.

One of the limitations of this study is that the code was developed based on limited data. Although the XGB model is robust and can generalize effectively to new data, however before any commercialization, it needs to be validated in the hospital and checked with additional new data to evaluate its performance and confirm generalization.

Moreover, our data came only from Iranian patients, and this single nationality could be considered a limitation. Also, we only considered patients who had ALS as a single diagnosis, and patients who had multiple neuropathic diseases at the same time, such as diabetes, rheumatoid arthritis, lymphoma, Lyme disease, etc., were excluded. Therefore in the future, data related to patients who suffer from multiple neuropathic diseases at the same time should be included in this code.

Conclusion

The high degree of complexity inherent in NCS and electromyography data makes it difficult to choose the best features to differentiate ALS from non-ALS conditions, especially radiculopathy. Our preliminary study using machine learning algorithms to address this problem has laid the groundwork for early diagnosis of this serious fatal disease. The results showed that the XGBoost algorithm showed the most accurate diagnosis include accuracy = 0.871, precision = 0.923, recall = 0.850, and f-1 score = 0.857 based on 35 features obtained from NCS.

Although the number of samples in our study was higher than many previous studies, the total number of patients was still small and a higher number of patients would lead to a more accurate conclusion.

Data availability

The findings of this study are derived from data that are available upon request from the corresponding author (FR).

Received: 17 August 2024; Accepted: 22 October 2025
Published online: 24 November 2025

References

1. Darabi, S. et al. Cerebrospinal fluid and blood exosomes as biomarkers for amyotrophic lateral sclerosis; a systematic review. *Diagn. Pathol.* **19**(1), 47 (2024).
2. Ketabforoush, A. et al. Sodium phenylbutyrate and tauroursodeoxycholic acid: a story of hope turned to disappointment in amyotrophic lateral sclerosis treatment. *Clin. Drug Investig.* **44**(7), 495–512 (2024).
3. Maksymowicz, S. & Siwek, T. Diagnostic odyssey in amyotrophic lateral sclerosis: diagnostic criteria and reality. *Neurol. Sci.* **45**(1), 191–196 (2024).
4. Jellinger, K.A., *The spectrum of behavioral disorders in amyotrophic lateral sclerosis: current view.* J. Neural Transmission (2024).
5. Bali, T. & Miller, T. M. Management of amyotrophic lateral sclerosis. *Mo Med* **110**(5), 417–421 (2013).
6. Trojsi, F., et al., *Resting state fMRI analysis of pseudobulbar affect in Amyotrophic Lateral Sclerosis (ALS): motor dysfunction of emotional expression.* Brain Imaging Behav (2022)
7. Carvalho, M, and Swash, M. Diagnosis and differential diagnosis of MND/ALS: IFCN handbook chapter. *Clin. Neurophysiol. Pract.* **9**, 27-38 (2024).
8. Fels, J. A. et al. Gene expression profiles in sporadic ALS fibroblasts define disease subtypes and the metabolic effects of the investigational drug EH301. *Hum. Mol. Genet.* **31**(20), 3458–3477 (2022).
9. Staats, K. A. et al. Blood-based biomarkers of inflammation in amyotrophic lateral sclerosis. *Mol. Neurodegener.* **17**(1), 11 (2022).
10. Sun, J. et al. Blood biomarkers and prognosis of amyotrophic lateral sclerosis. *Eur. J. Neurol.* **27**(11), 2125–2133 (2020).
11. Aydemir, D. and N.N. Ulusu, *Importance of the serum biochemical parameters as potential biomarkers for rapid diagnosis and evaluating preclinical stage of ALS.* Med. Hypotheses (2020)
12. Hashizume, A. et al. A functional scale for spinal and bulbar muscular atrophy: Cross-sectional and longitudinal study. *Neuromuscul. Disord.* **25**(7), 554–562 (2015).
13. Kimura, F. et al. Progression rate of ALSFRS-R at time of diagnosis predicts survival time in ALS. *Neurology* **66**(2), 265–267 (2006).
14. Goyal, N. A. et al. Misdiagnosis of amyotrophic lateral sclerosis in clinical practice in Europe and the USA: A patient chart review and physician survey. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **25**(1–2), 16–25 (2024).
15. Hill, J., Sanghani, N. & Li, Y. Features suggestive of coexisting amyotrophic lateral sclerosis in patients with spinal stenosis and influence of spinal decompression. *Cureus* **16**(1), e51587 (2024).
16. Semnani, M. R. et al. Evaluation of carotid Intima-Media Thickness (IMT) in amyotrophic lateral sclerosis disease using ultrasonography. *J. Clin. Neurosci.* **124**, 67–72 (2024).
17. Sathya, G. R. et al. F wave index: A diagnostic tool for peripheral neuropathy. *Indian J. Med. Res.* **145**(3), 353–357 (2017).
18. Fang, J. et al. Pattern differences of small hand muscle atrophy in amyotrophic lateral sclerosis and mimic disorders. *Chin. Med. J. (Engl)* **129**(7), 792–798 (2016).
19. Storti, B. et al. ALS mimics due to affection of the cervical spine: from common compressive myelopathy to rare CSF epidural collection. *Case Rep. Neurol.* **13**(1), 145–156 (2021).
20. Kwak, S. et al. Amyotrophic lateral sclerosis mimicking radiculopathy: a case series. *Nagoya J. Med. Sci.* **83**(4), 877–881 (2021).
21. Yamada, M., Furukawa, Y. & Hirohata, M. Amyotrophic lateral sclerosis: frequent complications by cervical spondylosis. *J. Orthop. Sci.* **8**(6), 878–881 (2003).
22. Carette, S. & Fehlings, M. G. Cervical radiculopathy. *N. Engl. J. Med.* **353**(4), 392–399 (2005).
23. Ellenberg, M. R., Honet, J. C. & Treanor, W. J. Cervical radiculopathy. *Arch. Phys. Med. Rehabil.* **75**(3), 342–352 (1994).
24. Rajagopalan, V., Chaitanya, K. G. & Pioro, E. P. Quantitative brain MRI metrics distinguish four different ALS phenotypes: A machine learning based study. *Diagnostics* **13**(9), 1521 (2023).
25. Behler, A., Müller, H. P., Ludolph, A. C. & Kassubek, J. Diffusion tensor imaging in amyotrophic lateral sclerosis: Machine learning for biomarker development. *Int. J. Mol. Sci.* **24**(3), 1911 (2023).
26. Ebrahimi Shah-abadi, M. et al. Recent advances and future directions in imaging of peripheral nervous system: A comprehensive review for therapeutics approach. *J. Adv. Med. Biomed. Res.* **31**(148), 415–431 (2023).
27. Guazzo, A. et al. Predicting clinical events characterizing the progression of amyotrophic lateral sclerosis via machine learning approaches using routine visits data: a feasibility study. *BMC Med. Inform. Decis Mak* **24**(Suppl 4), 318 (2024).
28. Chudzik, A., Śledzianowski, A. & Przybyszewski, A. W. Machine learning and digital biomarkers can detect early stages of neurodegenerative diseases. *Sensors* **24**(5), 1572 (2024).
29. Pancotti, C. et al. Deep learning methods to predict amyotrophic lateral sclerosis disease progression. *Sci. Rep.* **12**(1), 13738 (2022).
30. Tan, H. H. G. et al. MRI clustering reveals three ALS subtypes with unique neurodegeneration patterns. *Ann. Neurol.* **92**(6), 1030–1045 (2022).
31. Fan, J. et al. Differentiation between amyotrophic lateral sclerosis and mimics using quantitative analysis of fsciculation with muscle ultrasound. *Chin. Med. Sci. J.* **38**(4), 265–272 (2023).
32. Priyadarshini, T. S. & Hameed, M. A. Developing heart stroke prediction model using deep learning with combination of fixed row initial centroid method with Navie Bayes, Decision Tree, and Artificial Neural Network. *Measure. Sensors* **34**, 101237 (2024).
33. Das, A., Choudhury, D. & Sen, A. A collaborative empirical analysis on machine learning based disease prediction in health care system. *Int. J. Inf. Technol.* **16**(1), 261–270 (2024).
34. Tannemaat, M. R. et al. Distinguishing normal, neuropathic and myopathic EMG with an automated machine learning approach. *Clin. Neurophysiol.* **146**, 49–54 (2023).
35. Kalita, J. et al. Split hand index and ulnar to median ratio in Hirayama disease and amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Frontotemporal Degener* **18**(7–8), 598–603 (2017).
36. Menon, P. et al. Split-hand index for the diagnosis of amyotrophic lateral sclerosis. *Clin. Neurophysiol.* **124**(2), 410–416 (2013).
37. Souayah, N. et al. Electrodiagnostic profile of conduction slowing in amyotrophic lateral sclerosis. *Heliyon* **9**(8), e18400 (2023).
38. Stikvoort García, D. J. L. et al. Diagnostic accuracy of nerve excitability and compound muscle action potential scan derived biomarkers in amyotrophic lateral sclerosis. *European J. Neurol.* **30**(10), 3068–3078 (2023).
39. Antonescu, F. et al. Reassessing the Diagnostic utility of the split hand index in amyotrophic lateral sclerosis patients-the divide by zero problem. *Neurol. Int.* **14**(3), 707–715 (2022).
40. Sato, M. et al. Prolonged distal latency of the median motor nerve is associated with poor prognosis in amyotrophic lateral sclerosis. *Neurol. Res.* **43**(3), 191–198 (2021).

41. Hu, F. et al. Dissociated lower limb muscle involvement in amyotrophic lateral sclerosis and its differential diagnosis value. *Sci. Rep.* **9**(1), 17786 (2019).
42. Nishi, R. et al. Differential clinicopathologic features of EGPA-associated neuropathy with and without ANCA. *Neurology* **94**(16), e1726–e1737 (2020).
43. Daroff, R.B. and M.J. Aminoff, *Encyclopedia of the neurological sciences*. Academic press. (2014)
44. Li, J. et al. Feature selection: A data perspective. *ACM comput. Surv. (CSUR)* **50**(6), 1–45 (2017).
45. Louppe, G., et al., *Understanding variable importances in forests of randomized trees.* Advances in neural information processing systems (2013)
46. Altaf, I., Butt, M. A. & Zaman, M. Hard voting meta classifier for disease diagnosis using mean decrease in impurity for tree models. *Rev. Comput. Eng. Res.* **9**(2), 71–82 (2022).
47. Shastry, K. A. An ensemble nearest neighbor boosting technique for prediction of Parkinson's disease. *Healthcare Anal.* **3**, 100181 (2023).
48. Soltaninejad, S., A. Basu, and I. Cheng. *Automatic classification and monitoring of denovo parkinson's disease by learning demographic and clinical features*. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE. (2019)
49. Chen, T. and C. Guestrin. *Xgboost: A scalable tree boosting system*. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (2016)
50. Dhaliwal, S. S., Nahid, A.-A. & Abbas, R. Effective intrusion detection system using XGBoost. *Information* **9**(7), 149 (2018).
51. Sagi, O. & Rokach, L. Approximating XGBoost with an interpretable decision tree. *Inf. Sci.* **572**, 522–542 (2021).
52. Kiangala, S. K. & Wang, Z. An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment. *Mach. Learn. Appl.* **4**, 100024 (2021).
53. Mao, F. et al. An XGBoost-assisted evolutionary algorithm for expensive multiobjective optimization problems. *Inf. Sci.* **666**, 120449 (2024).
54. Sable, A. W. Median and ulnar nerves in the hand. *Phys. Med. Rehabil. Clin. N. Am.* **9**(4), 737–753 (1998).
55. Ryan, M. M. & Jones, H. R. Chapter 14 - Mononeuropathies. In *Neuromuscular Disorders of Infancy, Childhood, and Adolescence (Second Edition)* (eds Darras, B. T. et al.) 243–273 (Academic Press, 2015).
56. de Carvalho, M. & Swash, M. Nerve conduction studies in amyotrophic lateral sclerosis. *Muscle Nerve* **23**(3), 344–352 (2000).
57. Garcia, R. U. et al. Ulnar sensory-motor amplitude ratio: A new tool to differentiate ganglionopathy from polyneuropathy. *Arq. Neuropsiquiatr.* **71**(7), 465–469 (2013).
58. Aksoy, D. et al. Nerve conduction studies in the early diagnosis of amyotrophic lateral sclerosis and the importance of split-hand phenomenon. *Genel Tip Dergisi.* **32**(4), 451–454 (2022).
59. Vacchiano, V. et al. Comparative assessment of MScanFit MUNE and quantitative EMG in amyotrophic lateral sclerosis diagnosis: A prospective study. *Clin. Neurophysiol.* **166**, 66–73 (2024).
60. Fang, J. et al. A retrospective study of the characteristics and clinical significance of a-waves in amyotrophic lateral sclerosis. *Front. Neurol.* **8**, 515 (2017).
61. Zhang, J. et al. Quantitative evaluation of factors influencing the 3 Hz repetitive nerve stimulation test in patients with amyotrophic lateral sclerosis. *Muscle Nerve* **70**(2), 194–203 (2024).
62. Siddiqui, A. H. et al. Association of sensory nerve action potential amplitude and velocity with type 2 diabetic peripheral neuropathy. *Cureus* **15**(10), e46501 (2023).
63. Antunes, M. et al. A morphology-based feature set for automated Amyotrophic Lateral Sclerosis diagnosis on surface electromyography. *Biomed. Signal Process. Control* **79**, 104011 (2023).

Acknowledgements
Not Applicable.

Author contributions
Study design: F.R. Data preparation: A.A. and B.HA. Data analysis: F.R. and S.T. and M.A. Writing: F.R. and M.R.H. and S.T. and M.A. All authors approved the manuscript.

Funding
FR was supported by the IRAN University of Medical Sciences, Grant no. 1402–4-4–27191, and by the Shahid Beheshti University of Medical Sciences, Grant no. 43008282–0.

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval
This proposal was approved by Iran University of Medical Sciences ethics approval center with code number IR.IUMS.REC.1402.187. Due to the retrospective nature of this study, informed consent was waived by the Ethics Committee of Iran University of Medical Sciences.

Consent for publication
Not Applicable.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-25562-8.

Correspondence and requests for materials should be addressed to S.T. or F.R.