BioData Mining

RESEARCH

Open Access

Predictive modeling of ALS progression: an XGBoost approach using clinical features

Richa Gupta[1*], Mansi Bhandari[1], Anhad Grover[1], Taher Al-shehari[2], Mohammed Kadrie[2], Taha Alfakih[3] and Hussain Alsalman[4]

*Correspondence:
Richa Gupta
richagupta@jamiahamdard.ac.in

Full list of author information is available at the end of the article

Abstract

This research presents a predictive model aimed at estimating the progression of Amyotrophic Lateral Sclerosis (ALS) based on clinical features collected from a dataset of 50 patients. Important features included evaluations of speech, mobility, and respiratory function. We utilized an XGBoost regression model to forecast scores on the ALS Functional Rating Scale (ALSFRS-R), achieving a training mean squared error (MSE) of 0.1651 and a testing MSE of 0.0073, with R^2 values of 0.9800 for training and 0.9993 for testing. The model demonstrates high accuracy, providing a useful tool for clinicians to track disease progression and enhance patient management and treatment strategies.

Keywords Amyotrophic Lateral Sclerosis (ALS), ALS Functional Rating Scale (ALSFRS-R), Predictive modeling, Machine learning, Disease progression, XGBoost, Clinical features

Introduction

ALS is a progressive neurodegenerative disease marked by the loss of motor neurons, leading to muscle paralysis. Its causes remain unclear, and treatments aim to prolong life and manage symptoms. Clinical trials encounter challenges due to varying progression rates and lack of effective therapies. Diagnostic biomarkers and objective outcome measures are lacking, complicating ALS treatment [1, 2].

ALS can be familial (5% genetic) with SOD1 mutations, or sporadic (95%) with potential environmental and genetic factors. Key genes include C9orf72, SOD1, TARDBP, UBQLN2, and FUS. Pathogenesis involves inflammation, excitotoxicity, oxidative stress, and neurovascular issues, with "typical" ALS showing both upper and lower motor neuron loss, and atypical forms may not [3].

The ALS Functional Rating Scale-Revised (ALSFRS-R) is a validated tool with 12 items across four domains (bulbar, limb/trunk, respiratory, and interventions) used to assess functional deficits in ALS patients. ALSFRS-R is an updated version of the original ALSFRS, which consisted of 10 items and assessed respiratory function with a single question (Q10). In contrast, ALSFRS-R includes three distinct questions (Q10, Q11, Q12) to more thoroughly evaluate respiratory function, resulting in a maximum score of 48

points, compared to the 40-point maximum of the original ALSFRS. While ALSFRS-R is known for its reliability, some studies have noted inconsistencies in its measurement model across different contexts. Initially developed for clinical trials, ALSFRS-R is now widely used in research, clinical practice, self-assessments, and online questionnaires [4, 5].

The ALS Functional Rating Scale (ALSFRS-R) evaluates the functional abilities of ALS patients in areas such as speech, swallowing, mobility, and respiratory function. Each category is rated on a scale from 0 to 4, where a score of 4 represents normal function, and 0 indicates a complete loss of ability. Lower scores correspond to greater functional impairment. This tool is commonly used to track disease progression and assess patient status over time.

Pathophysiology

Oxidative stress in ALS leads to acetylated TDP-43 aggregates, which impair RNA binding and promote the accumulation of insoluble, hyperphosphorylated TDP-43 species that closely resemble pathological inclusions seen in ALS. These acetylated TDP-43 lesions have been observed in ALS patient spinal cords, linking aberrant TDP-43 acetylation to the pathogenesis of TDP-43 proteinopathy [6]. Aggregated TDP-43 sequesters specific microRNAs and mitochondrial proteins, further exacerbating mitochondrial dysfunction and oxidative stress, contributing to axonal degeneration and disease progression [7–9]. Mitochondrial dysfunction associated with ER-mitochondrial defects disrupts calcium balance, mitochondrial dynamics, and autophagy, which are essential to neuronal health in ALS [8]. Collectively, these mechanisms drive ALS pathology.

In this research, we draw upon the PRO-ACT database, utilizing datasets that include ALSFRS (R) scores, demographic information, disease duration, and clinical features to forecast ALS progression. By combining these varied datasets, our model provides a holistic perspective on the factors influencing ALS, allowing for more precise predictions of disease advancement. This methodology deepens our understanding of the clinical elements affecting ALS, thereby aiding in better patient management and outcomes.

Literature review

Pancotti et al. found "Onset Delta" as the top predictive feature for ALS progression. The FFNN+CNN model showed slightly lower error but also lower correlation in ALSFRS slope prediction. Shapley values improved model interpretability, and fast progressors were associated with significantly shorter survival [1].

Vieira et al. developed ML models to predict ALSFRS-R scores using voice and accelerometer data. A CNN achieved an AUC of 0.86 for speech-related scores, and accelerometer models had AUCs of 0.70–0.75 for limb functions. The models, showing strong correlations with self-reported assessments, offer potential as objective ALS evaluation tools and provide insights into ALS progression and edaravone treatment effects [2].

Müller et al. used an RNN to predict breathing decline in ALS patients, based on ALSFRS-R question 10 scores. SHAP values identified key factors affecting breathing capacity, including speech and swallowing, highlighting the importance of interpretable models for personalized ALS care [10].

The CompALS project used dynamic Bayesian networks to predict ALS progression with ITIS and IT datasets. The models accurately predicted disease progression and

survival, assessing risk factors like onset site and FVC, providing valuable insights for clinical decision-making and personalized care [11].

A multinational study created a predictive model for ALS outcomes using data from 1936 patients across 14 European centers. The Royston-Parmar model, which included predictors like onset type and C9orf72 mutation, achieved a concordance statistic of 0.78 and was externally validated, aiding in patient stratification and personalized management [12].

Imamura et al. developed a CNN-based model using iPSC-derived motor neurons, achieving 90% accuracy and an AUC of 0.97. The VGG-16 network outperformed human experts and traditional models, with Grad-CAM visualizations highlighting neuron changes, showing CNNs' potential for improving ALS prediction accuracy [13].

Abdul Jabbar et al. developed ML models with PRO-ACT data from 5030 ALS patients, using XGBoost and BLSTM, achieving AUROC values of 0.570 to 0.748. The models identified 21 key predictors and showed potential to reduce Phase II/III clinical trial sizes by 18.3%, enhancing ALS research and trial design [14].

Albert A. Taylor et al. developed predictive models for ALS progression using PRO-ACT data from over 10,700 patients. The random forest model, with 13 predictors, showed superior accuracy and stability compared to pre-slope and GLM models, highlighting its clinical potential for improving ALS progression predictions [15].

Alberto Tena's study analyzed speech in 45 ALS patients and 18 controls, using PCA to extract features like jitter and pitch. Supervised classification models effectively distinguished between ALS and control participants, highlighting acoustic analysis's potential for diagnosing ALS-related speech changes [16].

Jahandideh et al. developed a GBM model to predict forced vital capacity (FVC) in ALS patients using PRO-ACT data. The model showed reliable accuracy and was validated internally and externally, with key features including "Baseline FVC" and "Days since baseline [17]. Below is a Table 1 comparing the related works presented in the literature review section:

Methodology

In this study, we apply an XGBoost regression model to predict ALS progression, utilizing clinical features such as speech, mobility, and respiratory function. The model was trained and fine-tuned for high efficiency, yielding accurate predictions on the ALS Functional Rating Scale (ALSFRS-R).

In clinical practice, there's rising interest in using machine learning algorithms like XGBoost for developing AI models. XGBoost, a boosting algorithm, is notable for refining errors from pre-existing models [18]. The Mean Squared Error (MSE) for regression is mathematically represented by Eq. (1):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\mathrm{y_i} - \widehat{\mathrm{y}}_\mathrm{i})^2 \tag{1}$$

Datasets

This research utilized data sourced from the PRO-ACT database, which contains extensive information on ALS patients. The datasets obtained included the ALS Functional

Table 1 Comparative analysis of related works discussed in the literature review

Reference Number	Research Focus	Methodology	Key Findings	Limitations
[1]	ALSFRS-R factor analysis	EFA and CFA on ALSFRS-R	Four-factor model with cross-loading items fits best	Generalizability limited to Dutch patients
[2]	ALS severity measurement	ML models using voice and accelerometer data	Accurate prediction of bulbar and limb ALSFRS-R scores	ML models limited to dataset characteristics
[10]	ALS progression prediction	RNNs with LSTMs; SHAP for explainability	Low MSE (< 0.02); key features identified	Explainability challenges; limited outcome scope
[11]	ALS progression modeling	Dynamic Bayesian Networks (DBNs)	Predicts functional impairment and survival with high accuracy	Dependent on initial visit data
[12]	ALS survival prediction	Royston-Parmar model, external validation	Identified 8 predictors, c-statistic: 0.78, distinct survival groups	Requires medical doctor application
[13]	ALS diagnosis support	Deep learning, iPSCs, CNN	AUC: 0.97 for classifying healthy vs. ALS	Requires further prospective research
[14]	ALS progression prediction	XGBoost, BLSTM, PRO-ACT database	AUROC: 0.570–0.748, 21 key predictors identified	Confidence levels vary
[15]	ALS disease progression prediction	Random Forest (RF), GLM, pre-slope models	RF Model superior in predicting disease progression, validated on clinical and trial data	Potential overfitting in non-RF models
[16]	Diagnosis of bulbar involvement in ALS	Automated voice analysis, SVM	Automated model (accuracy 95.8%) outperforms human diagnosis	Limited to Spanish vowel analysis
[17]	Prediction of vital capacity in ALS	Gradient boosting machine (GBM) with PRO-ACT dataset	ALS-VC model predicts vital capacity with RMSD of ~ 0.534, similar across datasets	Limited external validation data

Rating Scale (ALSFRS) scores, forced vital capacity (FVC), demographic details, onset history, and mortality records.

Machine learning methodology

- **Programming Language**: Python
- **Machine Learning Libraries**:
  - **XGBoost**: Used for implementing the gradient boosting algorithm to enhance prediction accuracy.
  - **Scikit-learn**: Employed for model evaluation and validation, including train-test splitting and cross-validation.
- **Data Analysis Libraries**:
  - **Pandas**: Utilized for data manipulation and analysis, allowing for efficient handling of the dataset.
  - **NumPy**: Used for numerical computations and to facilitate array operations.
  - **Matplotlib and Seaborn**: Employed for data visualization, providing insights through various plots and graphs.
  - **Missingno**: Used for visualizing missing data patterns in the dataset.

Data integration

The separate datasets were combined using 'subject_id' as the merging key, with an outer join applied to retain all records, resulting in a consolidated dataset referred to as 'merged_data.

Data preparation

- Managing Missing Data

For numeric features with missing entries, the mean was used for imputation, whereas categorical variables were transformed using label encoding. Records with excessive missing values in essential variables were removed to ensure the integrity of the dataset.

- Feature Selection

Selected features for modeling comprised critical clinical indicators such as disease duration, ALSFRS scores, and respiratory measurements. A correlation matrix was created to analyze the interrelationships among these variables.

Feature engineering

To enhance the predictive capacity of the model, feature engineering was employed. One critical feature introduced was the Progression Rate, which quantifies the rate of disease deterioration. It was calculated as:

$$\text{Progression Rate} = \frac{48 - \text{ALSFRS} - \text{R Total}}{\text{Disease Duration}}$$

This derived feature provided valuable insight into the patient's condition and was a key predictor in the regression model. Additionally, interaction terms between certain features, such as combining respiratory function with mobility scores, were explored to capture complex relationships that could influence ALS progression. All features were scaled using z-score normalization to ensure consistency across the dataset and improve model performance.

Model development

- Choice of Model

The XGBoost regression model was chosen for its strong performance in regression tasks. The model was configured to use 50 estimators with a maximum depth of 6.

- Training and Validation

The dataset was allocated into a training set (90%) and a testing set (10%), utilizing a random state to ensure reproducibility. Cross-validation techniques were applied to assess the model's performance, resulting in a mean squared error (MSE) of 4.6225 (±4.7155).

Model assessment

The model's performance was assessed through mean squared error (MSE) and R^2 scores. The training dataset yielded an MSE of 0.1651 and a training R^2 of 0.9800, demonstrating a high level of fit.

The workflow for the XGBoost machine learning model is presented, encompassing data preprocessing, model training, and evaluation, thereby offering a thorough understanding of its methodology (see Fig. 1).

Result and discussion

Preprocessing

The preprocessing phase involved several critical steps to ensure data quality and readiness for analysis. Date columns, such as Date of Birth, were converted to datetime format for accurate calculations of age and disease duration. Missing values in numeric columns were filled using the mean, while categorical variables were transformed through label encoding. Important features, including ALSFRS-R scores and respiratory metrics, were selected based on their clinical significance. Additionally, new features, such as the Progression Rate, were created to reflect the rate of disease progression. Finally, all numeric features underwent z-score normalization to improve model convergence and overall performance.

Model performance

The predictive model exhibited strong performance metrics, with a cross-validation mean squared error (MSE) of 4.6225 (±4.7155). The training dataset resulted in an MSE of 0.1651, while the test dataset demonstrated a markedly lower MSE of 0.0073. The R^2 values further highlighted the model's efficacy, with a training R^2 of 0.9800 and a test R^2 of 0.9993, indicating an exceptional fit to the data.

Figures 2 and 3 shows the actual versus predicted ALSFRS-R Total values, while the accompanying heatmap and swarm plot provide further insights into the model's performance and feature relationships.

Feature importance

Analysis of feature importance indicated that specific clinical indicators, such as disease duration and respiratory measurements, were pivotal in predicting ALS progression. These results are consistent with existing research emphasizing the significance of functional measures in the progression of ALS. As shown in Fig. 4, the feature importance highlights the key predictors utilized in the model.

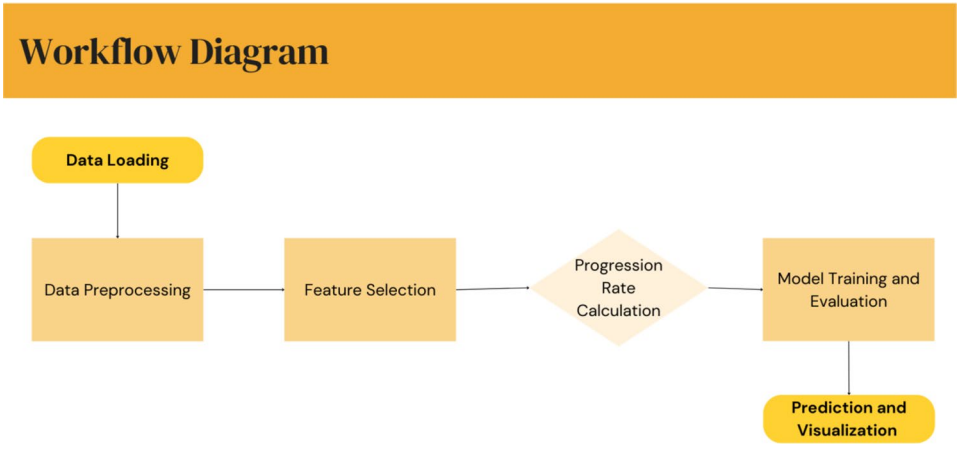

Fig. 1 Detailed workflow of model

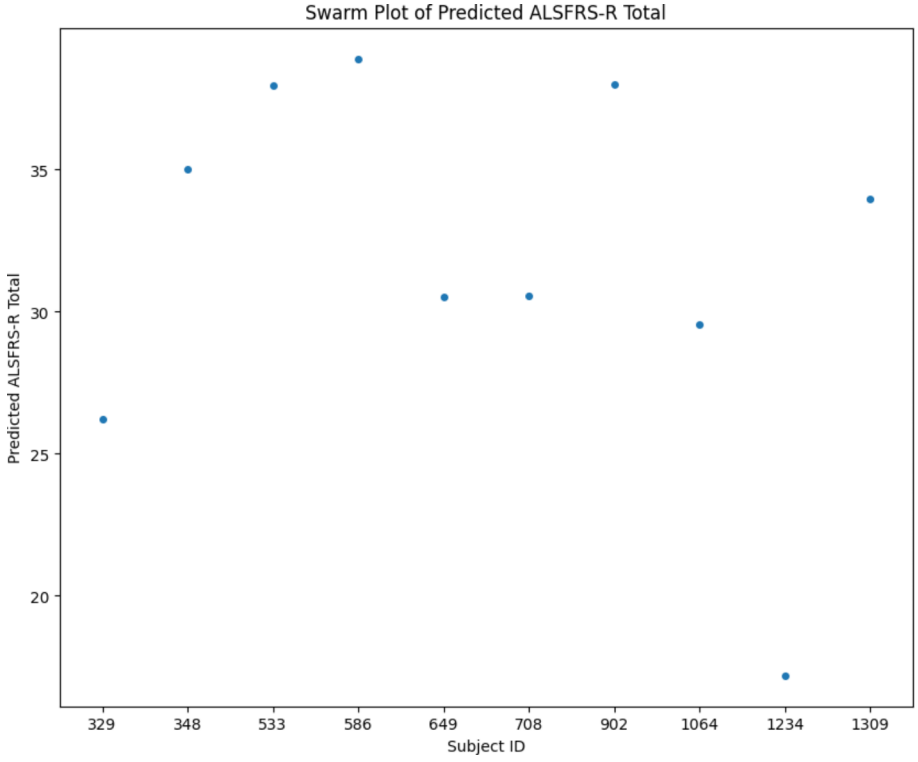

Fig. 2 Swarm plot of predicted ALSFRS-R total

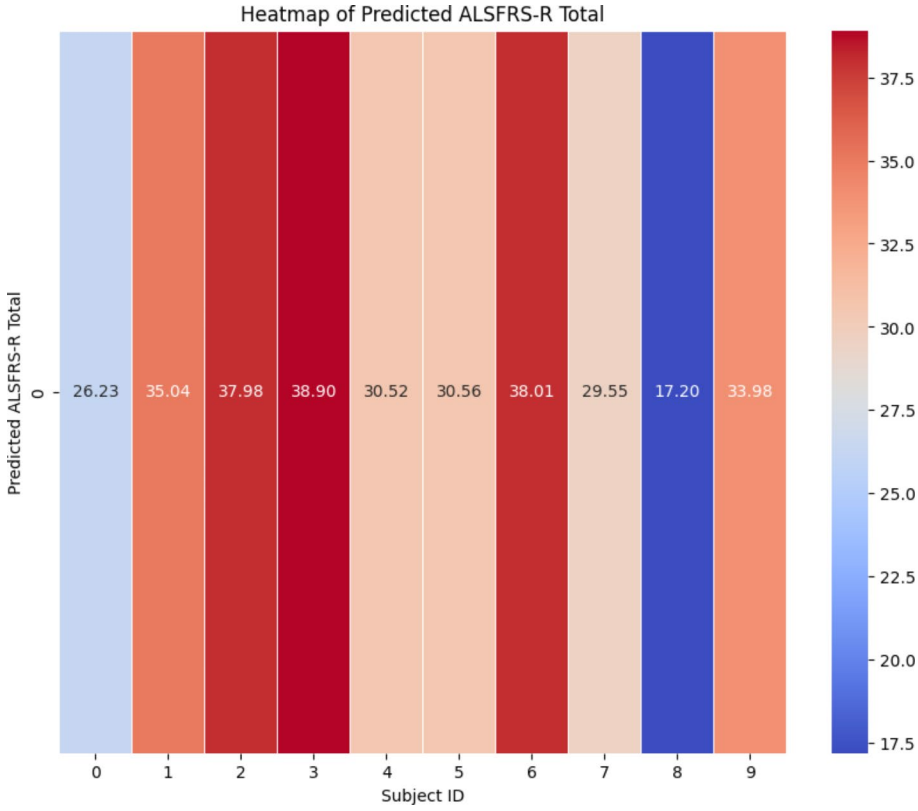

Fig. 3 Heatmap of predicted ALSFR-S total

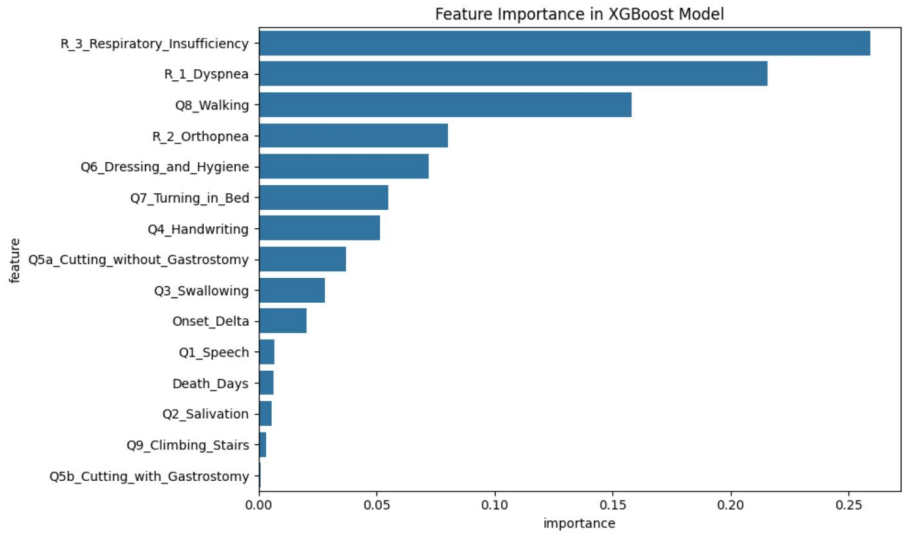

Fig. 4 Feature importance of model (Rank-wise)

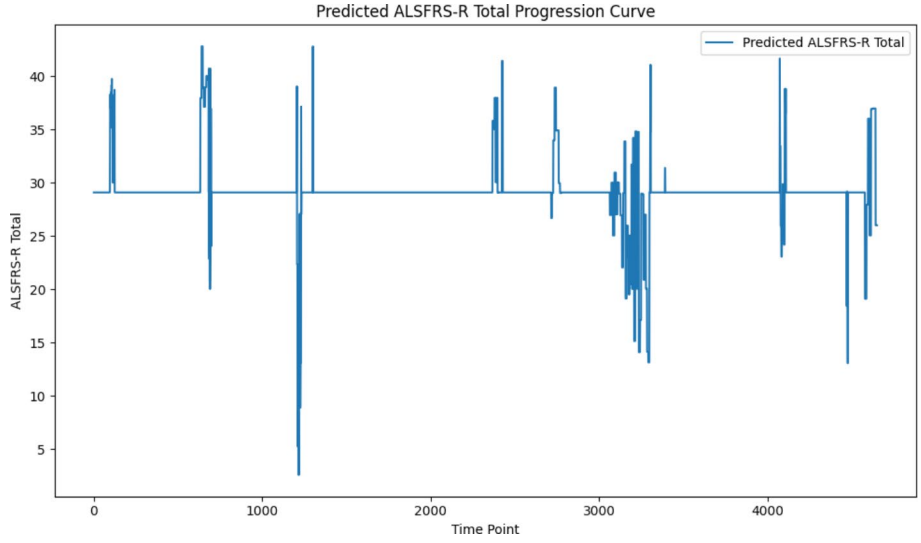

Fig. 5 Predicted ALSFRS-R total progression curve

Predicted progression

The model's predictions showed a steady decline in ALSFRS scores over time, offering valuable insights into expected patient trajectories. Visualizations, including progression curves and heatmaps, effectively illustrated the predicted decline across various subject IDs, aiding in a better understanding and potential planning for patient care. The progression curve depicted in Fig. 5 illustrates the predicted trajectory of ALSFRS-R Total scores over time. The predicted progression curve for 10 subject IDs is illustrated in Fig. 6.

Below is the Table 2 comparing the results of related works with the current study.

The results of this study highlight the potential of machine learning models to predict ALS progression based on clinical features. The low MSE and high R^2 values suggest that the model is robust and capable of generalizing to new data, making it a useful tool for clinicians. The findings underscore the importance of disease duration and respiratory

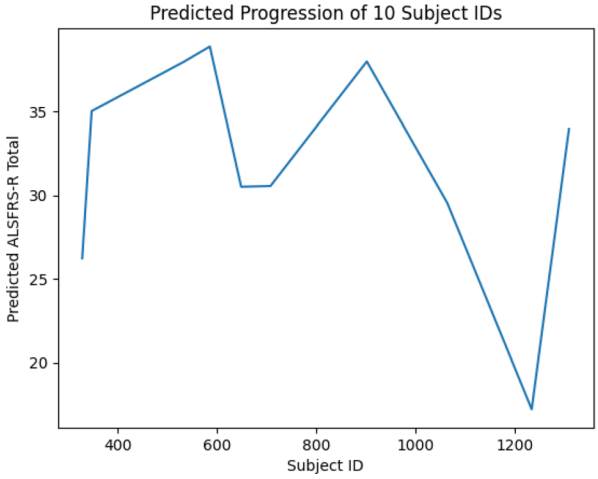

Fig. 6 Predicted ALSFRS-R total progression curve for selected subjects

Table 2 Comparison of results from related works with the current study

Reference Number	Previous Work	Current Work
[1]	ALSFRS-R factor analysis using EFA and CFA; best fit: four-factor model	Utilizes XGBoost for predicting ALS progression based on clinical features and demographic data.
[2]	ML models using voice and accelerometer data for ALS severity	Focuses solely on predicting ALS progression using XGBoost based on clinical assessments.
[10]	ALS progression prediction with RNNs and SHAP; low MSE (< 0.02)	Achieves an XGBoost MSE of 0.0073 and R^2 of 0.9993, showcasing high predictive accuracy.
[11]	ALS progression modeling with Dynamic Bayesian Networks; high accuracy	Implements XGBoost for ALS progression prediction and evaluates its performance.
[12]	ALS survival prediction using Royston-Parmar model; c-statistic: 0.78	Concentrates specifically on predicting ALS progression rather than survival outcomes.
[13]	ALS diagnosis support with deep learning and CNN; AUC: 0.97	Emphasizes ALS progression prediction using XGBoost rather than diagnostic support.
[14]	ALS progression prediction with XGBoost and BLSTM; AUROC: 0.570–0.748	XGBoost demonstrates superior performance with an R^2 value of 0.9993 in predicting ALSFRS scores.
[15]	ALS disease progression prediction with RF, GLM, and pre-slope models; RF superior	Compares various models, highlighting XGBoost as the primary method for ALS progression prediction.
[16]	Diagnosis of bulbar involvement using automated voice analysis and SVM; accuracy 95.8%	Focuses exclusively on ALS progression prediction without diagnosing bulbar involvement.
[17]	Prediction of vital capacity in ALS using GBM; RMSD ~ 0.534	Utilizes XGBoost to predict ALS progression, achieving low MSE values for high accuracy.

function as key predictors, which is in line with prior research emphasizing the need to monitor these parameters in patient management.

Moreover, the visualizations produced from this analysis offer an intuitive representation of patient progression, which could prove beneficial for both healthcare providers and caregivers in making informed decisions. However, while the model shows considerable promise, further validation with larger and more diverse datasets is necessary to confirm its effectiveness across various populations.

Future research should also consider incorporating additional variables, such as genetic markers and lifestyle factors, to enhance predictive accuracy. By continuously refining these models, we can improve personalized care strategies and potentially facilitate earlier interventions in the disease process.

Conclusion

In summary, this study demonstrates the effectiveness of machine learning approaches in predicting ALS progression, highlighting the significance of specific clinical features. The model serves as a promising tool for improving patient management and care.

Future work

Looking ahead, we plan to further enhance our predictive model by leveraging advanced techniques, including more refined feature engineering, model optimization, and exploring additional machine learning approaches. Our continued efforts are focused on improving the precision and reliability of ALS progression predictions, with the ultimate goal of advancing the understanding and treatment of this complex neurodegenerative condition.

Acknowledgements
This research was supported by the Researchers Supporting Project number (RSP2024R244), King Saud University, Riyadh, Saudi Arabia.

Author contributions
R.G. (Dr. Richa Gupta) and M.B. (Mansi Bhandari) designed the study and wrote the main manuscript text. A.G. (Anhad Grover) conducted data collection and performed computational analyses. T.A. (Taher Al-shehari) provided methodological insights and reviewed the manuscript. M.K. (Mohammed Kadrie) assisted with result interpretation and drafted sections of the manuscript. T.Af. (Taha Alfakih) was involved in data acquisition, analysis, and discussion of findings. H.A. (Hussain Alsalman) prepared data visualizations and assisted in manuscript preparation. All authors reviewed and approved the final manuscript.

Funding
This research was supported by the Researchers Supporting Project number (RSP2024R244), King Saud University, Riyadh, Saudi Arabia.

Data availability
The data supporting the results of this study are available from the PRO-ACT database. Access to the PRO-ACT database can be requested via the following link: https://nctu.partners.org/ProACT. Further data supporting the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Computer Science, King Saud University, Riyadh, Saudi Arabia
[2]Computer Skills, Department of Self-Development Skill, Common First Year Deanship, King Saud University, Riyadhi 11362, Saudi Arabia
[3]Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[4]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Received: 28 August 2024 / Accepted: 11 October 2024
Published online: 02 December 2024

References

1. Pancotti C, Birolo G, Rollo C, Sanavia T, Di Camillo B, Manera U, Chiò A, Fariselli P. Deep learning methods to predict amyotrophic lateral sclerosis disease progression. Sci Rep. 2022;12:Art13738.
2. Vieira FG, Venugopalan S, Premasiri AS, McNally M, Jansen A, McCloskey K, Brenner MP, Perrin S. A machine-learning based objective measure for ALS disease severity. Npj Digit Med. 2022;5:Art45.
3. Wijesekera LC, Leigh PN. Amyotrophic lateral sclerosis. Orphanet J Rare Dis, 4, Art. 3, 2009.

4. Maier A, Boentert M, Reilich P, Witzel S, Petri S, Großkreutz J, Metelmann M, Lingor P, Cordts I, Dorst J, Zeller D, Günther R, Hagenacker T, Grehl T, Spittel S, Schuster J, Ludolph A, Meyer T. ALSFRS-R-SE: an adapted, annotated, and self-explanatory version of the revised amyotrophic lateral sclerosis functional rating scale. Neurol Res Pract, 4, Art. 60, 2022.
5. Bakker LA, Schröder CD, van Es MA, Westers P, Visser-Meily JMA, van den Berg LH. Assessment of the factorial validity and reliability of the ALSFRS-R: a revision of its measurement model. J Neurol. Jun. 2017;264:1413–20.
6. Zuo X, Zhou J, Li Y, Wu K, Chen Z, Luo Z, Zhang X, Liang Y, Esteban MA, Zhou Y, Fu X-D. TDP-43 aggregation induced by oxidative stress causes global mitochondrial imbalance in ALS. Nat Struct Mol Biol. 2021;28:132–42.
7. Cohen TJ, Hwang AW, Restrepo CR, Yuan C-X, Trojanowski JQ, Lee VMY. An acetylation switch controls TDP-43 function and aggregation propensity. Nat Commun, 6, Art. 5845, 2015.
8. Bernard-Marissal N, Chrast R, Schneider BL. Endoplasmic reticulum and mitochondria in diseases of motor and sensory neurons: a broken relationship? Cell Death Dis., vol. 9, Art. no. 333, 2018.
9. Izumikawa K, Nobe Y, Yoshikawa H, Ishikawa H, Miura Y, Nakayama H, Nonaka T, Hasegawa M, Egawa N, Inoue H, Nishikawa K, Yamano K, Simpson RJ, Taoka M, Yamauchi Y, Isobe T, Takahashi N. TDP-43 stabilises the processing intermediates of mitochondrial transcripts. Sci Rep, 7, Art. 7709, 2017.
10. Müller M, Gromicho M, de Carvalho M, Madeira SC. Explainable models of disease progression in ALS: learning from longitudinal clinical data with recurrent neural networks and deep model explanation. Comput Methods Programs Biomed. 2021;1:Art100018.
11. Tavazzi E, Daberdaku S, Zandonà A, Vasta R, Nefussy B, Lunetta C, Mora G, Mandrioli J, Grisan E, Tarlarini C, Calvo A, Moglia C, Drory V, Gotkine M, Chiò A. B. Di Camillo, and For the Piemonte, Valle d'Aosta Register for ALS (PARALS) and the Emilia Romagna Registry for ALS (ERRALS), Predicting functional impairment trajectories in amyotrophic lateral sclerosis: a probabilistic, multifactorial model of disease progression, *J. Neurol.*, vol. 269, no. 7, pp. 3858–3878, Mar. 2022, https://doi.org/10.1007/s00415-022-11022-0
12. Westeneng H-J, Debray TPA, Visser AE, van Eijk RPA, Rooney JPK, Calvo A, et al. Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. Lancet Neurol. May 2018;17(5):423–33.
13. Imamura K, Yada Y, Izumi Y, Morita M, Kawata A, Arisato T, Nagahashi A, Enami T, Tsukita K, Kawakami H, Nakagawa M, Takahashi R, Inoue H et al. Jun., Prediction model of amyotrophic lateral sclerosis by deep learning with patient induced pluripotent stem cells, Ann. Neurol., vol. 89, no. 6, pp. 1226–1233, 2021.
14. Abdul Jabbar MAD, Guo L, Nag S, Guo Y, Simmons Z, Pioro EP, Ramasamy S, Yeo CJJ. Predicting amyotrophic lateral sclerosis (ALS) progression with machine learning, Amyotroph. Lateral Scler. Frontotemporal Degener., vol. 25, no. 3–4, pp. 242–255, Dec. 2023.
15. Taylor AA, Fournier C, Polak M, Wang L, Zach N, Keymer M, Glass JD, Ennist DL, The Pooled Resource Open-Access ALS Clinical Trials Consortium. Predicting disease progression in amyotrophic lateral sclerosis. Ann Clin Transl Neurol. Nov. 2016;3(11):866–75.
16. Tena A, Claria F, Solsona F, Meister E, Povedano M. Detection of bulbar involvement in patients with amyotrophic lateral sclerosis by machine learning voice analysis: Diagnostic decision support development study, JMIR Med. Inform., vol. 9, no. 3, Mar. 2021.
17. Jahandideh S, Taylor AA, Beaulieu D, Keymer M, Meng L, Bian A, Atassi N, Andrews J, Ennist DL. Longitudinal modeling to predict vital capacity in amyotrophic lateral sclerosis, Amyotroph. Lateral Scler. Frontotemporal Degener., vol. 19, no. 3–4, pp. 294–302, Dec. 2017.
18. Ogunleye A, Wang Q-G. XGBoost model for chronic kidney disease diagnosis. IEEE/ACM Trans Comput Biol Bioinf. 2020;17(6):1808–16. Nov.-Dec.

Publisher's note