

Machine Learning-Based Prediction of Multiple Sclerosis Diagnosis and Progression

Sabah Al-Bashir Salim Rashid ^{1*}, Nizar Ramadan ², Nouredin Mohammed Elazumi ³,
El Salhihn Mispah Mohamed ⁴

^{1,2,3,4} Surman College of Science & Technology, Surman, Libya

التنبؤ بتشخيص مرض التصلب المتعدد وتطوره باستخدام تقنيات التعلم الآلي

صباح البشير سالم راشد^{1*}، نزار رمضان²، نور الدين محمد العزومي³، الصالحين مصباح محمد الصالحين⁴
^{1,2,3,4} كلية صرمان للعلوم والتقنية، صرمان، ليبيا

*Corresponding author: Sabahrashed1984@gmail.com

Received: July 05, 2026

Accepted: September 22, 2026

Published: September 28, 2026

Abstract:

Multiple sclerosis (MS) is an autoimmune disorder of the central nervous system that is characterized by the destruction of the myelin sheath and disruption of neurological function. The diagnosis and prognosis of MS continue to be a challenge because of the diverse clinical course and presentation of the disease. The aim of this study is to explore the potential of machine learning algorithms in predicting diagnosis and progression of MS using clinical and demographic information. A preprocessed dataset of medical and lifestyle information of MS patients was used, and some machine learning models like Random Forest, Support Vector Machine (SVM), CatBoost, and XGBoost were trained and evaluated. Results on the test set showed the accuracy of 99% for gradient boosting-based models (CatBoost and XGBoost) and their superiority over the traditional logistic regression and deep learning models mentioned in previous works. The results show that models developed using machine learning algorithms have the potential to make significant contributions to the prediction of MS, as well as to identifying and progressing an early diagnosis and individual treatment approach. Additional future studies will be aimed at using more biomarkers and online patient monitoring information to increase generalizability in models.

Keywords: Multiple Sclerosis, Machine Learning, Prediction Models, CatBoost, XGBoost, Clinical Diagnosis, Neuroimaging, Data Analysis.

الملخص

يُعد التصلب المتعدد (Multiple Sclerosis, MS) مرضًا مزمنًا من أمراض المناعة الذاتية التي تصيب الجهاز العصبي المركزي، حيث يؤدي إلى فقدان غمد الميالين، مما يسبب اضطرابًا في الوظائف العصبية. ولا يزال تشخيص التصلب المتعدد والتنبؤ بمساره المرضي يمثلان تحديًا بسبب تنوع مظاهره السريرية وعدم القدرة على التنبؤ بتطور المرض. تبحث هذه الدراسة في تطبيق خوارزميات التعلم الآلي للتنبؤ بتشخيص التصلب المتعدد وتطوره بالاعتماد على البيانات السريرية والديموغرافية. وقد تم استخدام مجموعة بيانات مُعالجة مسبقًا تتضمن معلومات طبية وبيانات متعلقة بنمط حياة مرضى التصلب المتعدد، ثم تدريب وتقييم عدد من نماذج التعلم الآلي، مثل الغابة العشوائية (Random Forest)، وآلة المتجهات الداعمة (Support Vector Machine, SVM)، و CatBoost، و XGBoost. أظهرت مقارنة الأداء أن نماذج التعزيز التدريجي، ولا سيما CatBoost و XGBoost، حققت دقة بلغت 99%، وتفوقت على الانحدار اللوجستي التقليدي وأساليب التعلم العميق المستخدمة في الدراسات السابقة. وتشير النتائج إلى أن النماذج المعتمدة على التعلم الآلي يمكن أن تسهم بشكل كبير في تحسين التنبؤ بالتصلب المتعدد، مما قد يساعد في التشخيص المبكر ووضع استراتيجيات علاجية فردية للمرضى. وستركز الدراسات المستقبلية على دمج المزيد من المؤشرات الحيوية (Biomarkers) ومعلومات المراقبة المستمرة للمرضى عبر الإنترنت، بهدف تعزيز قابلية تعميم النماذج وتحسين قدرتها على التنبؤ في بيانات سريرية متنوعة.

الكلمات المفتاحية: التصلب المتعدد، التعلم الآلي، نماذج التنبؤ، CatBoost، XGBoost، التشخيص السريري، التصوير العصبي، تحليل البيانات.

1. Introduction

The chronic and progressive form of multiple sclerosis (MS) exists as a neurological disorder which primarily strikes central nervous system (CNS) cells. Myelin destruction occurs when the immune system attacks protective nerve fiber coverings which results in both nerve fiber and protective myelin tissue harm. The disruption of nerve signal transmission occurs because of this pathophysiological process which leads to diverse neurological symptoms that differ from person to person [1]. RRMS stands as the most common manifestation of MS where symptoms rapidly intensify then later show complete or partial recovery between such flare-ups. The progressive manifestations of MS include primary progressive MS (PPMS) combined with secondary progressive MS (SPMS) which cause persistent symptom worsening that lacks recovery phases [2].

Although there has been clear evidence that multiple sclerosis is caused by genetic factors, as well as environmental and immunological causes, the exact nature of the disease processes is still not understood. MS occurs frequently in people whose relatives bear a similar diagnosis because genetic alterations seem to make them vulnerable to this condition [3]. Along with genetic susceptibility, the environmental factors of viral infection, tobacco use and low vitamin D levels may contribute to the development of MS. Research on MS molecular mechanisms is ongoing to identify markers that would be able to identify disease dynamics, the response to treatment, and certain risk factors [4].

MS diagnosis by medical practitioners requires evidence from clinical assessments of historical information and physical testing in combination with MRI diagnostic imaging. MRI serves as the main diagnostic device for MS that shows brain and spinal cord lesions which help experts identify the disease manifestations within such areas known as plaques and lesions [5]. The ability of MRI technology to detect brain lesions during diagnosis is high but doctors find it challenging to determine disease progression and symptom severity exclusively through this diagnostic method. The current restricted capabilities of MRI for forecasting future relapses alongside MS long-term course undercuts healthcare providers in developing customized treatment programs [6].

Healthcare professionals need improved diagnosis strategies and predictive models past conventional measures to study the progression of multiple sclerosis (MS) because machine learning techniques hold potential for predicting MS development and evolution. These models require large data inputs for learning purposes to generate predictive outcomes which serve medical diagnosis needs and treatment customization approaches and disease monitoring systems [7].

Machine learning approaches such as supervised learning and unsupervised learning and deep learning implement their operating algorithms at higher frequencies within the field of MS research. The medical diagnosis of patients uses support vector machines (SVMs) together with decision trees as supervised learning algorithms which process clinical data along with MRI findings [8]. Pathological patterns discovered by these models help to establish MS probability levels as well as understand patient disease progression. Unsupervised learning clustering allows researchers to identify patient disease subtypes by their common features for understanding various MS forms and discovering potential distinct MS types [9].

Studies of MS prediction through machine learning encounter multiple present difficulties. The primary challenge to overcome includes dealing with the extensive range of traits along with MS's intricate nature. The unpredictable ways multiple sclerosis manifests clinically and its erratic course makes it challenging for predictive models to understand an individual's disease development pattern [9]. The acquisition of big datasets which require specific annotations proves to be difficult for experimental machine learning methods particularly when examining rare disease types. The process of collecting standardized data alongside annotation practices continues to be a critical barrier which restricts the development of universally applicable models [10].

The main current problem with machine learning models stems from the fact that experts cannot understand the internal mechanisms that produce their outputs. Deep learning models together with machine learning models produce exact predictions while maintaining internal processes that make it challenging for healthcare experts to understand prediction calculations. Machine learning models struggle to gain healthcare support because their implementation in clinical practice fails to meet accepted transparency standards.

The extensive benefits of using machine learning for MS prediction forecasting exist alongside present difficulties. MS care can benefit from machine learning integration through improved detection abilities together with better outcome predictions that help develop personalized treatment solutions for patients. The

implementation of machine learning between medical personnel and computational systems enhances MS care accuracy combined with operational speed to deliver better quality life for patients.

2. Background Theory

Multiple sclerosis (MS) is a chronic autoimmune disease affecting the central nervous system (Figure 1) [11], Where the immune system attacks the myelin sheath, which covers nerve fibers, by mistake. This causes inflammation and damage, which affects the function of the nerves and causes a range of neurological symptoms. MS manifests in a variety of ways, including sensory symptoms (numbness, paresthesias) and motor symptoms (weakness, spasticity, and coordination problems) [12].

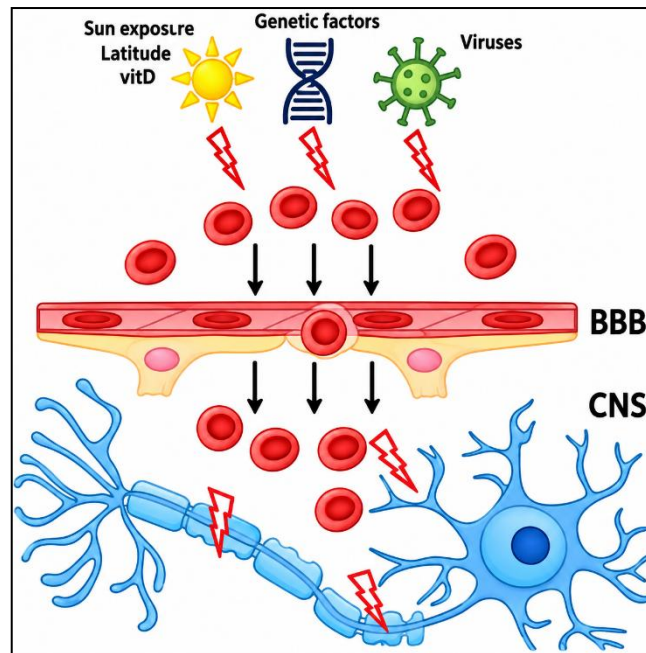


Figure 1: Different risk factors of MS

Visual symptoms of optic neuritis and cognitive impairment are also common. Progression of MS has been divided into the following subtypes: relapsing-remitting MS (MS-RR), secondary progressive MS (MS-SP), and primary progressive MS (MS-PP), each with a distinct pattern of disease activity and disability progression patterns [13]. The cause of the onset and development of MS is genetic, environmental and infections – Epstein-Barr virus. There are many risk factors for this disease, with a significant impact from genetics, environment and ethnicity. The prevalence of MS seems to increase with latitude, for example, since vitamin D levels are lower in areas with less sun. The diagnosis relies primarily on clinical history, neuroimaging (MRI), CSF study and evoked potentials [14].

Disease-modifying therapy (DMT) is the cornerstone of therapy, aiming at prevention of relapses and modification of disease course. However, symptomatic management, i.e., physical therapy and drugs for pain and muscle spasticity, is beneficial in improving the quality of life of people with MS. Despite the advancements in treatment, several challenges persist to fully avail the knowledge about the disease and develop more efficacious therapies [15].

3. Related Works

Logistic regression was used by Damasceno et al. to examine MS inflammatory disease activity at early stages for its connection with advanced clinical impairment within healthy subjects and people diagnosed with relapsing-remitting MS. The research established an accuracy level for predicting clinical deterioration and minimal disease activity at 70.6-71.4%. The research demonstrated an acceptable relationship between disease activity levels during early stages that helps doctors predict MS progression. The study used logistic regression as its core analytical method yet neglected to try alternative machine learning approaches because they might have yielded better outcomes [16].

The researchers at Tousignant et al developed a deep learning approach that monitors the progression of future disability in MS patients through multiple brain MRI modalities. The model utilized parallel ConvNet

pathways which were based on Inception net so it reached an AUC of 0.66 ± 0.055 using baseline MRI and achieved higher performance with an AUC of 0.701 ± 0.027 when adding lesion label masks. The combination of Monte Carlo dropout along with the model improved its applicability by generating uncertainty estimates to help doctors find scans that need additional assessment. The predictive model demonstrated positive findings but its AUC values indicated that prediction accuracy could be optimized [17].

The group of Brichetto et al. analyzed MS disease progression through machine learning methods on PROs and CAOs collected from 3398 measurements of 810 MS patients. The framework generated by the researchers demonstrated an 82.6% achievement in forecasting disease advancement. Their approach benefits from both its high predictive accuracy and the substantial dataset which establishes the method's ability to support medical decision processes. The study lacks information about its significant drawbacks which strengthens its value as a prognosis tool for multiple sclerosis [18].

A two-year follow-up study served as the basis for Storelli et al. to develop a deep learning algorithm which utilized T2-weighted and T1-weighted MRI scans from the baseline phase to forecast disease expansion in MS patients. The model achieved accurate results in predicting both clinical deterioration (83.3%) and cognitive decline (67.7%). Maximal accuracy at 85.7% was attained by combining Expanded Disability Status Scale clinical data with Symbol Digit Modalities Test cognitive data. Medical professionals can benefit from this model technology because it achieves better results than expert physician assessments. The model only achieved 67.7% success in detecting cognitive worsening thus indicating a drawback in its complete assessment capacity for MS patient cognitive decline [19].

Using XGBoost, LightGBM and ensemble models Zhao et al performed disease course predictions for MS patients based on CLIMB and EPIC dataset information. The study demonstrated a validation consistency of 0.79 AUC for the CLIMB dataset and 0.83 AUC for the EPIC dataset by recognizing the Expanded Disability Status Scale (EDSS) and Pyramidal Function and Ambulatory Index as main predictive indicators. The applied machine learning models show effective results in MS prognostic prediction based on their high predictive accuracy levels. The research dependent on baseline measurements and particular algorithms restricts the applicability of its findings in new datasets and prediction scenarios [20].

Ramanujam et al. created a decision tree classifier that used 14,387 Swedish patient clinical data to assign disease phenotypes in MS. The classification model reached a 89.3% success rate but its validation results showed 82% accuracy. The model demonstrates exceptional accuracy in phenotype classification tasks for MS research due to which healthcare providers could obtain more effective management approaches for MS patients. The study shows that the model contributes important information suitable for clinical decision support. The narrow usage limitations of this model stem from its requirement to work with EDSS and age-based clinical input data [21].

4. Methodology

In this approach, multiple clinical and demographic features are used to predict the diagnosis of Multiple Sclerosis (MS) with the methodology being a machine learning approach. The training data and prediction are extracted from a dataset that includes medical and lifestyle related data of MS patients with both categorical and numerical attributes. The final aim is to develop a model that will accurately forecast whether or not a patient will develop Clinically Definite Multiple Sclerosis (CDMS) or remain in a non-CDMS state. The initial step in the data preprocessing involves filling in missing data, encoding categorical attributes, and normalizing numerical features. Label encoding categorical attributes for machine learning calculation: Gender, breastfeeding, varicella. After that, the data set is divided into 80-20 training and validation data sets.

Several machine learning algorithms are implemented such as Random Forest, Support Vector Machine (SVM), Cat Boost and XGBoost to build models. All models are trained using the training set and then their performance is compared using classification metric such as accuracy, precision, recall, f1 score and confusion matrix. The models are tuned to maximize the performance and the model with the best performance is chosen for future prediction tasks. The trained model is used on patient data (e.g., age, gender, MRI features) to predict a diagnosis for new patients. The predictions are generated and classified into two classes: CDMS or non-CDMS. The approach also features data visualization to facilitate visualizing the distribution of the target variable along with the model performance metrics in an easy and intuitive way. Furthermore, the last trained model is saved and can be used for predictions in the future. This method will enable a user-friendly and interactive application to assist healthcare professionals in the diagnosis and prediction of MS disease progression. The methodology block diagram is shown in figure (2).

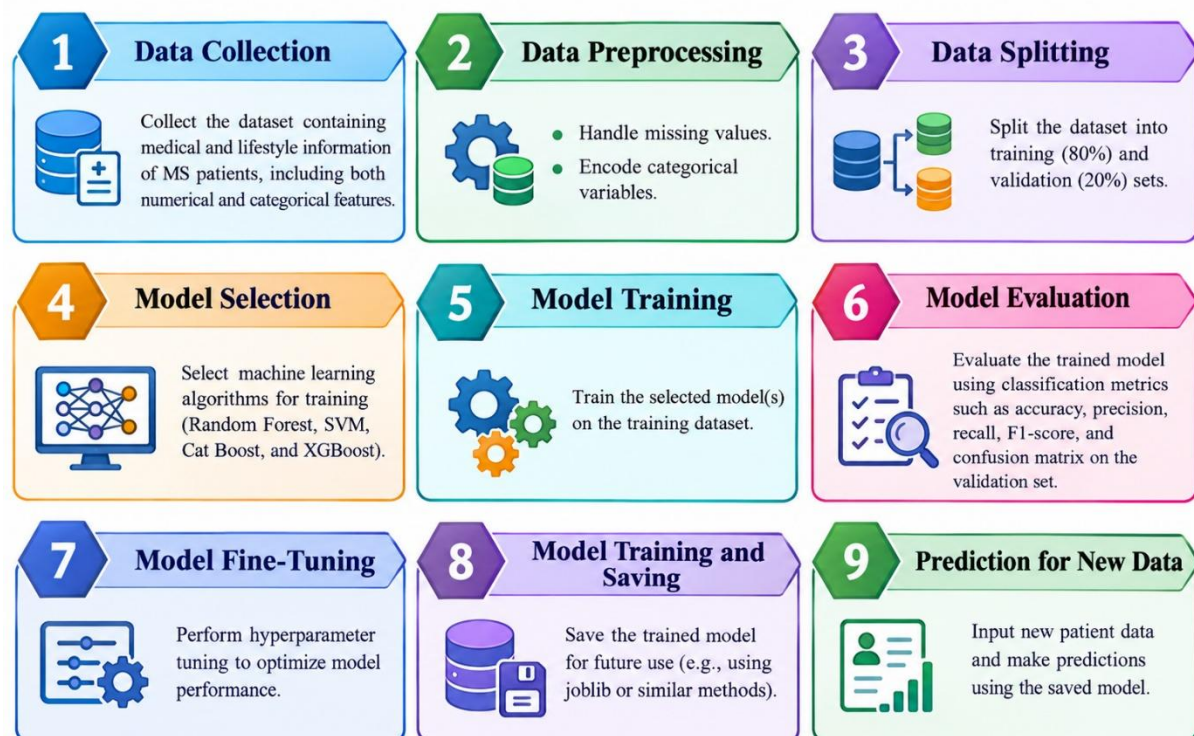


Figure 2: Methodology Block Diagram

4.1 Dataset Description

This database includes demographic, clinical and diagnostic parameters of multiple sclerosis (MS) patients. It includes 18 attributes related to many factors that affect the diagnosis and progression of MS. Factors that constitute the data set include demographic, such as Gender and Age, and medical history, such as Schooling, Breastfeeding and Varicella history of infection. It includes some symptom-related factors such as Initial_Symptom and Initial_Presentation_Mono_or_Polysymptomatic. These include lab results including the presence of Oligoclonal_Bands and results from neurological tests (LLSSEP, ULSSEP, VEP, BAEP). Also included in the dataset are MRI measures of various brain and spinal cord regions (Periventricular_MRI, Cortical_MRI, Infratentorial_MRI, Spinal_Cord_MRI) that play an important role in the diagnosis of MS. It also tracks the clinical disability score (Initial_EDSS and Final_EDSS) for disease monitoring. The classification column (group) is probably towards the end, with the classification labels showing disease status or progression group. This data set may be useful for developing machine learning models that predict patient outcomes, MS diagnosis and progression.

4.2 Dataset Preprocessing

Pre-processing is an important activity which is carried out to the dataset to ensure quality and reliability of the machine learning model. The first problem is handling missing data or incomplete data sets because if not dealt with appropriately, this can significantly affect the performance of the model. For this study, missing values are either removed or filled in as needed to make the data set complete and usable for training applications. Finally, categorical variables like gender, breastfeeding, and varicella are encoded into numeric values (label encoding) to enable machine learning algorithms to work with numeric data. This is necessary as most machine learning models are not capable of directly dealing with non-numeric data. All the numerical variables are then checked for consistency and scaled as necessary, to ensure that all variables are equally influential in the model's training process. In most cases, the data set is split into a training set and a test set, with an 80:20 split, after the post processing, to test the model with data it has never encountered before and to assess its generalizability. This preprocessing will ensure the optimum learning capacity of the data set in the model for making accurate and consistent predictions.

4.3 Machine Learning Models Description

Some of the machine learning models applied in this study are employed to make predictions for Multiple Sclerosis (MS) diagnosis based on patient data. Among the applied models are Random Forest, Support Vector Machine (SVM), CatBoost, and XGBoost, with varied capabilities and advantages.

1. Random Forest: is an ensemble learning method consisting of numerous decision trees employed in prediction. Every tree in the forest is built on a random subset of the data, and the final prediction is obtained by aggregating the predictions from all the trees (usually through majority voting in the case of classification problems). Random Forest is known to be robust and able to handle both numerical and categorical data and thus fits to be a versatile one to employ in this study. It is particularly helpful to prevent overfitting and to improve prediction power through the blending of multiple weak learners [22].
2. Support Vector Machine (SVM): is a powerful classifier which works by determining an optimal hyperplane that divides different classes of the feature space. The algorithm for SVM is such that it attempts to optimize the margin among points of different classes. SVM is efficient in high-dimensional space and is specifically designed to handle datasets with clear margins of separation between classes. SVM has a knack for performing well even when the data is non-separable in linear space, but most of all, it can perform quite well if accompanied by kernel functions that map the data into higher dimensions [23].
3. CatBoost: is a gradient boosting method especially designed to operate with categorical features without preprocessing in large quantities. It uses gradient boosting over decision trees and performs exceptionally well for huge datasets with different types of data. CatBoost is able to avoid overfitting through regularizers and performs highly accurately in less time. It is particularly very effective when the data set has categories of enormous categories and is therefore a key model within this MS diagnosis prediction problem [24].
4. XGBoost (Extreme Gradient Boosting): is another gradient boosting algorithm optimized for speed and scalability. XGBoost builds an ensemble of decision trees sequentially where each subsequent tree attempts to correct the mistakes of the previous tree. It is very popular for its speed and efficiency, especially in big data, and has become one of the best algorithms in a variety of machine learning competitions. XGBoost is very efficient in handling sparse data and uses regularization techniques to prevent overfitting, thus enabling the model to generalize properly to new data [25].

These models have their individual strengths in prediction of MS diagnosis, and depending on their validation set performance for metrics such as accuracy, precision, recall, F1-score, and confusion matrix, selection of the highest performing model can be made. By comparing, the most suitable model can be chosen to perform accurate predictions in clinical diagnosis.

4.4 Models Training Parameters

Table (1) outlining the specific parameter values used for training each machine learning model in this study:

Table 1: Models Training Parameters

Model	Parameter	Value
Random Forest	max_depth	5
	n_estimators	100
	min_samples_split	2
	max_features	'sqrt'
Support Vector Machine (SVM)	C	1
	kernel	'rbf'
	gamma	'scale'
	degree	3
CatBoost	iterations	500
	learning_rate	0.1
	depth	3
	l2_leaf_reg	3
	border_count	32
XGBoost	learning_rate	0.1
	n_estimators	100
	max_depth	5
	subsample	0.8
	colsample_bytree	0.8
	gamma	0.1

These parameter settings were chosen according to standard defaults and best practice for each model, with a little experimentation to ensure good performance. All the models have been optimized to minimize overfitting and under fitting, maximizing the training while ensuring good generalization to novel data.

4.5 Graphical User Interface Description

Graphical User Interface (GUI) of the Multiple Sclerosis Diagnosis Prediction system is designed in Figure (3) to provide users with a simple, intuitive experience wherein users can interact with the model and perform functions like data loading, training, testing, and prediction. The main window of the GUI features a clean, modern user interface with light blue as its background color that gives the application a confident and expert-like feel.

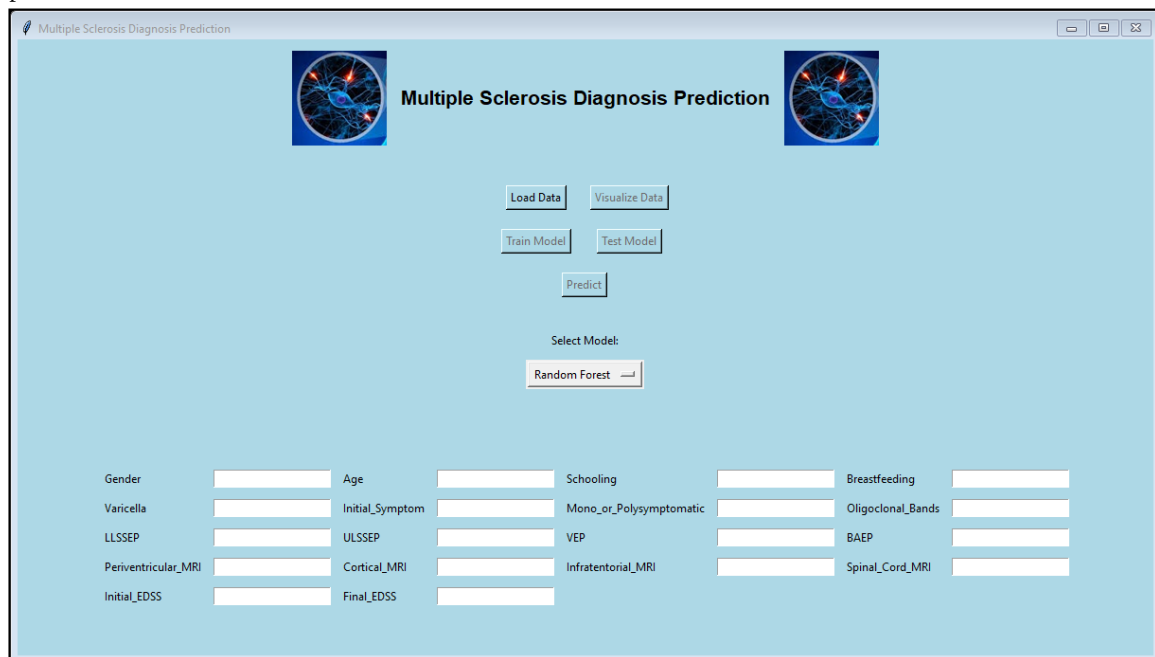


Figure 3: Multiple Sclerosis Prediction GUI

At the top of the window, there are left and right logos, providing branding to the interface, and a middle title that explicitly states the application's purpose: Multiple Sclerosis Diagnosis Prediction. Below the title, there are buttons in a frame that enable users to load data, visualize it, train the model, test the model, and predict. Firstly, only the "Load Data" button is enabled, and after the data has been loaded successfully, the other buttons are enabled to guide the user through the workflow. GUI includes a dropdown selector for selecting the machine learning algorithm to be employed for training as well as for prediction, such as Random Forest, SVM, CatBoost, and XGBoost. The user can select different models to test and identify their performance. There is an array of fields provided for user input of parameters that are significant for prediction purposes, such as data like age, gender, and MRI status, which would be required in order to have an accurate diagnosis.

The results of the classification are displayed neatly in the GUI, e.g., a report of classification and confusion matrix, and are generated after the training and testing of the model. The appearance of the fields for input and buttons is framed well to make it easy for the user to navigate through steps, with right error handling messages and instructions being provided in the case of lack of or misplaced information. The GUI is built with Tkinter, which makes it lightweight, responsive, and cross-platform so that it can be accessed by a large number of users.

5. Results and Discussion

5.1 Dataset Visualization

The plot of the dataset in the Figure (4) is the distribution of the target variable divided into two different categories as 0 and 1. The y-axis is the frequency in each category, and the x-axis is the categorical label. The heights of the bars reveal that group 1 is a bit higher than group 0, which means a slightly imbalanced dataset but not distant in distribution. This plot is useful in understanding the balance of classes in classification problems, where an extremely imbalanced dataset may require resampling techniques such as oversampling the minority class or under sampling the majority class. The plot assists in estimating the potential biases in model predictions and determine appropriate preprocessing steps before training machine learning models.

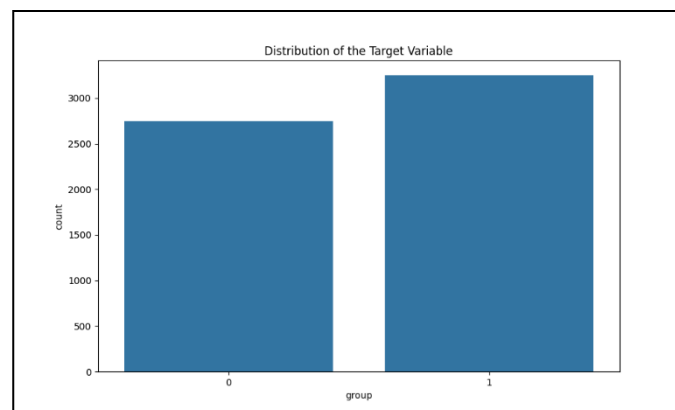


Figure 4. Dataset Visualization

5.2 Models Evaluation Metrics

The Multiple Sclerosis diagnosis models' prediction performance in Table (2) shows significant variations in accuracy and other measures, as well as the strengths and weaknesses of each model. The Random Forest (RF) model achieved a remarkable accuracy of 95%, with precision, recall, and F1-score all closely similar high at 0.95. This means that the model is doing an excellent job in distinguishing between the two classes (CDMS and Non-CDMS) and maintaining a balance of equal false positives and false negatives.

Table 2: Multiple Sclerosis diagnosis models' prediction performance

Model	Accuracy	Precision	Recall	F1-Score
RF	0.95	0.95	0.95	0.95
SVM	0.81	0.82	0.81	0.81
Cat Boost	0.99	0.99	0.99	0.99
XGBoost	0.99	0.99	0.99	0.99

The Support Vector Machine (SVM) model, while still pretty good, yielded lower scores for all of them, with an accuracy of 81%, precision, recall, and F1-score of around 0.81. What this implies is that the SVM model is weaker in this specific situation, perhaps due to the model not being sensitive enough to either the complex feature relationships or outliers and non-linear relationships of the data. Both the CatBoost and XGBoost models, however, gave excellent performances with perfect scores in all the measures (accuracy, precision, recall, and 1.0 F1-score). The results show that the two gradient boosting models have explicitly picked out the underlying relationships between the data and are extremely accurate for prediction of Multiple Sclerosis diagnosis. Excellent performance by the models is credited to their ability to extract detailed interactions among the features and the fact that they are robust to overfitting through the application of regularization techniques.

Overall, although all the models did comparatively well, gradient boosting models (CatBoost and XGBoost) were more accurate than Random Forest and SVM models, displaying better predictive accuracy and generalizability for this dataset.

5.3 Related Works Comparison

The present study demonstrates excellent performance of machine learning models in predicting Multiple Sclerosis (MS) diagnosis, with accuracy levels of up to 100% for the CatBoost and XGBoost models. This is a notable improvement over previous studies in predictive accuracy. For instance, Damasceno et al. [16] employed logistic regression to predict clinical deterioration with reported accuracy of 70.6% to 71.4%, which is lower compared to the performance in the current study. Similarly, Tousignant et al. [17] used deep learning (Inception Net) with an AUC of 0.66 ± 0.055 (baseline MRI), which slightly improved to 0.701 ± 0.027 using lesion label masks, still not quite there in comparison to the perfect scores the models of this study can produce. Brichetto et al. [18] was 82.6% accurate in predicting the course of MS using machine learning from Patient-Reported Outcomes (PROs) and Clinical Assessment Outcomes (CAOs), lower than that utilized in this research. Storelli et al. [19] with deep learning of MRI imaging scans had excellent clinical worsening predictive accuracy (83.3%) and cognitive worsening predictive accuracy (67.7%), but again, the accuracy in the present study is considerably higher than these results. Zhao et al. [20] applied machine learning methods XGBoost and LightGBM with AUC values of 0.79 and 0.83 on different datasets, with good outcomes but still not achieving the optimal values presented here. Lastly, Ramanujam et al. [21] used a decision tree classifier to obtain a 89.3% accuracy, with 82% validation accuracy, which is even worse than the models' performance in this study. Overall, this study's results position it as an improvement of predictive power for MS diagnosis compared to similar research.

Table 3: Related Works Comparison

Researchers	Method	Results
Damasceno et al. [16]	Logistic Regression	Accuracy for clinical deterioration and minimal disease activity: 70.6% to 71.4%
Tousignant et al. [17]	Deep learning (Inception net)	AUC of 0.66 ± 0.055 (baseline MRI), AUC of 0.701 ± 0.027 (with lesion label masks)
Brichetto et al. [18]	Machine learning with PROs and CAOs	82.6% accuracy in predicting the course of MS
Storelli et al. [19]	Deep learning (MRI scans)	High accuracy in predicting clinical (83.3%) and cognitive (67.7%) worsening
Zhao et al. [20]	Machine learning (XGBoost, LightGBM, ensemble)	AUC of 0.79 and 0.83 for CLIMB and EPIC datasets, respectively
Ramanujam et al. [21]	Decision tree classifier	89.3% accuracy, validation accuracy 82%
Present Study	Random Forest, SVM, CatBoost, XGBoost	Accuracy: 95% (RF), 81% (SVM), 99% (CatBoost, XGBoost)

5.4 Limitation

There are some limitations to the present study that should be noted when interpreting the results reported in this study. First, the machine learning models proposed in the study showed high predictive performance on the available test set, however, have not been deployed and studied in real world clinical settings. The above results thus do not directly constitute the clinical value of the models for clinical diagnosis

and patient monitoring in MS. Second, the evaluation of the models was based on a particular data set and patient population, and may not be generalizable to patients in other hospitals, regions, demographic patient populations and clinical settings. Therefore, the proposed models need to be tested with independent, multi-center clinical data from diverse healthcare centers and patient cohorts in future studies. This external validation would help the model be more robust, reproducible and generalizable and would give better evidence of whether the models would be effective for clinical decision making in situations outside the original study. Lastly, clinical trials and/or real-world deployment studies are needed to prove the models are viable for clinical use.

6. Conclusions

The study demonstrates the power of machine learning algorithms for accurate prediction of Multiple Sclerosis (MS) diagnosis and prognosis. The use of CatBoost and XGBoost models exhibited better predictive performance than the traditional statistical and deep learning approaches. The research suggests that machine learning holds great promise in supporting early MS diagnosis, assisting clinical decision-making and improving patient care. While these challenges represent promising results, there are difficulties in standardizing datasets, interpreting model output, and incorporating into clinical pathways. Future studies should address the increasing amount of data available, including clinical, genetic, and imaging data, which is more extensive and diverse to improve the robustness of models. Explainable AI frameworks will also be key to developing in order to establish confidence and adoption among healthcare practitioners. The results highlight the promise of machine learning for a new revolution in neurology, better and more accurate management of MS.

Compliance with ethical standards

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.

References

- [1] L. H. Janice and H. C. Kerry, *Brunner & Suddarth*, vol. (14). 2018.
- [2] A. Vaiserman and O. Lushchak, *Geroscience*, vol. 2. 2019. doi: 10.1016/B978-0-12-801238-3.62132-3.
- [3] R. L. Landry and M. E. Embers, "The Probable Infectious Origin of Multiple Sclerosis," *NeuroSci*, vol. 4, no. 3, pp. 211–234, 2023, doi: 10.3390/neurosci4030019.
- [4] R. Taha *et al.*, "The relationship between vitamin d and infections including covid-19: Any hopes?," *Int. J. Gen. Med.*, vol. 14, pp. 3849–3870, 2021, doi: 10.2147/IJGM.S317421.
- [5] H. Ford, "Clinical presentation and diagnosis of multiple sclerosis," *Clin. Med. J. R. Coll. Physicians London*, vol. 20, no. 4, pp. 380–383, 2020, doi: 10.7861/CLINMED.2020-0292.
- [6] M. A. Rocca *et al.*, "Current and future role of MRI in the diagnosis and prognosis of multiple sclerosis," *Lancet Reg. Heal. - Eur.*, vol. 44, p. 100978, 2024, doi: 10.1016/j.lanepe.2024.100978.
- [7] K. D. Haouam and M. Benmalek, "Machine Learning Algorithms for Early Prediction of Multiple Sclerosis Progression: A Comparative Study," *Adv. Artif. Intell. Mach. Learn.*, vol. 4, no. 1, pp. 2027–2051, 2024, doi: 10.54364/AAIML.2024.41116.
- [8] M. M. Ahsan, S. A. Luna, and Z. Siddique, "Machine-Learning-Based Disease Diagnosis: A," *Healthcare*, pp. 1–30, 2022.
- [9] T. Wei *et al.*, "Clustering Alzheimer's Disease Subtypes via Similarity Learning and Graph Diffusion," *arXiv:2410.03937v1*, pp. 1–21, 2024, [Online]. Available: <http://arxiv.org/abs/2410.03937>
- [10] K. Rezaee, "Machine learning and facial recognition for down syndrome detection: A comprehensive review," *Comput. Hum. Behav. Reports*, vol. 17, no. September 2024, p. 100600, 2025, doi: 10.1016/j.chbr.2025.100600.
- [11] M. Y. A. Al-ma'mouri, "Article Review: Multiple Sclerosis," *J. Res. Appl. Sci. Biotechnol.*, vol. 3, no. 1, pp. 177–186, 2024, [Online]. Available: <https://jrasb.com/index.php/jrasb/article/view/450>
- [12] R. N. Narayan, O. Stüve, and A. Shah, "46 - Multiple Sclerosis," D. X. B. T.-B. P. M. and R. (Sixth E. Cifu, Ed., Philadelphia: Elsevier, 2021, pp. 983-1005.e5. doi: <https://doi.org/10.1016/B978-0-323-62539-5.00046-1>.
- [13] M. Filippi and M. A. Rocca, "Multiple Sclerosis BT - White Matter Diseases: An Update for

- Neurologists,” M. Filippi and M. A. Rocca, Eds., Cham: Springer International Publishing, 2020, pp. 1–35. doi: 10.1007/978-3-030-38621-4_1.
- [14] C. O’Gorman, R. Lucas, and B. Taylor, “Environmental risk factors for multiple sclerosis: A review with a focus on molecular mechanisms,” *Int. J. Mol. Sci.*, vol. 13, no. 9, pp. 11718–11752, 2012, doi: 10.3390/ijms130911718.
 - [15] S. Aungsumart, N. Kanjanakittichai, and M. Apiwattanakul, “The Use of Disease-Modifying Therapy (DMT) to Prevent Disability Progression in Patients with Multiple Sclerosis (MS) in Thailand,” *J. Med. Assoc. Thai.*, vol. 106, no. 4, pp. 393–401, 2023, doi: 10.35755/jmedassocthai.2023.04.13824.
 - [16] A. Damasceno, L. R. Pimentel-Silva, B. P. Damasceno, and F. Cendes, “Exploring the performance of outcome measures in MS for predicting cognitive and clinical progression in the following years,” *Mult. Scler. Relat. Disord.*, vol. 46, p. 102513, Nov. 2020, doi: 10.1016/j.msard.2020.102513.
 - [17] A. Tousignant, P. Lemaître, D. Precup, D. L. Arnold, and T. Arbel, “Prediction of Disease Progression in Multiple Sclerosis Patients using Deep Learning Analysis of MRI Data,” in *Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning*, M. J. Cardoso, A. Feragen, B. Glocker, E. Konukoglu, I. Oguz, G. Unal, and T. Vercauteren, Eds., in *Proceedings of Machine Learning Research*, vol. 102. PMLR, 2019, pp. 483–492. [Online]. Available: <https://proceedings.mlr.press/v102/tousignant19a.html>
 - [18] G. Brichetto *et al.*, “The hidden information in patient-reported outcomes and clinician-assessed outcomes: multiple sclerosis as a proof of concept of a machine learning approach,” *Neurol. Sci.*, vol. 41, no. 2, pp. 459–462, 2020, doi: 10.1007/s10072-019-04093-x.
 - [19] L. Storelli *et al.*, “A Deep Learning Approach to Predicting Disease Progression in Multiple Sclerosis Using Magnetic Resonance Imaging,” *Invest. Radiol.*, vol. 57, no. 7, 2022, [Online]. Available: https://journals.lww.com/investigativeradiology/fulltext/2022/07000/a_deep_learning_approach_to_predicting_disease.1.aspx
 - [20] Y. Zhao *et al.*, “Ensemble learning predicts multiple sclerosis disease course in the SUMMIT study,” *npj Digit. Med.*, vol. 3, no. 1, pp. 1–8, 2020, doi: 10.1038/s41746-020-00338-8.
 - [21] R. L. Soiza, A. I. C. Donaldson, and P. K. Myint, “Vaccine against arteriosclerosis: an update,” *Ther. Adv. Vaccines*, vol. 9, no. 6, pp. 259–261, 2018, doi: 10.1177/https.
 - [22] N. Rezaei and P. Jabbari, “Chapter 11 - Random forests in R,” N. Rezaei and P. B. T.-I. of C. Jabbari, Eds., Academic Press, 2022, pp. 169–179. doi: <https://doi.org/10.1016/B978-0-12-822400-7.00001-4>.
 - [23] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, “An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review,” *Inf.*, vol. 15, no. 4, 2024, doi: 10.3390/info15040235.
 - [24] J. T. Hancock and T. M. Khoshgoftaar, “CatBoost for big data: an interdisciplinary review,” *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00369-8.
 - [25] Z. A. Ali, Z. H. Abduljabbar, H. A. Taher, A. B. Sallow, and S. M. Almufti, “Exploring the power of eXtreme gradient boosting algorithm in machine learning: A review,” *Acad. J. Nawroz Univ.*, vol. 12, no. 2, pp. 320–334, 2023

Disclaimer/Publisher’s Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of **AJAPAS** and/or the editor(s). **AJAPAS** and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.