Artificial Intelligence & Robotics Development Journal
Volume 6, Issue 6, pp 428-437, January 2026,https://doi.org/10.52098/airdj.20266169
ISSN: 2788-9696 Received: 14/11/2025 Revised: 23/01/2026 Accepted: 2/2/2026

Early Classification of Clinically Definite Multiple Sclerosis Using Machine Learning on Multimodal Diagnostic Features: A Random Forest-Based Clinical Decision Model

[1]Maryam Abu-Shauer & [2]Randa Abdelkareem
[1]Faculty of Neurology, Psychiatry and Physical Rehabilitation, Kyiv Medical niversity, Kyiv, Ukraine
* Corresponding author: Maryam Abu-Shauer[1] , maryamshauer@gmail.com
[2]Independent Researcher, Sohar, Sultanate of Oman.
randaabusham@gmail.com

Abstract

Early and precise diagnosis of Clinically Definite Multiple Sclerosis (CDMS) continues to be a significant issue in neurology, especially during the progression from initial symptoms to a conclusive diagnosis. This study introduces a machine learning-driven predictive model that utilizes multimodal diagnostic data—comprising MRI results, somatosensory evoked potentials (LLSSEP and ULSSEP), visual and auditory evoked potentials (VEP, BAEP), and clinical-demographic variables—to categorize patients as CDMS or non-CDMS. A Random Forest classifier was developed and verified using a dataset of 273 actual patients. The model had a very high predictive performance, with an area under the ROC curve (AUC) of more than 0.90 and very good precision-recall metrics. Feature importance analysis indicated that spinal cord MRI, oligoclonal bands, and lower limb SSEPs are the most significant predictors of CDMS. These results correspond with established biomarkers of MS development, illustrating the model's clinical significance. The suggested framework provides understandable, highly accurate decision assistance that could help neurologists diagnose MS early and figure out how likely someone is to have it. This research advances the emerging domain of AI-driven neurodiagnostics and underscores the efficacy of integrating electrophysiological, imaging, and history patient data for precise classification of neuroinflammatory illnesses. Future validation in multi-center cohorts may facilitate its implementation in clinical decision systems.

Keywords*:* multiple sclerosis; machine learning; early classification; multimodal diagnostic features; random forest

1. Introduction

The Multiple Sclerosis (MS) is a chronic longitudinal neurological disease that is attributed to the immune system sparking inflammation, demyelination and neurodegeneration of the central nervous system (Yousif, et al., 2025). The ability to distinguish between Clinically Definite Multiple Sclerosis (CDMS) and non-CDMS early and properly is significant to treat the disease well and help improve patient outcomes. However, the various attributes of the symptoms of MS and the limitations of the traditional methods of diagnosis often pose significant challenges to timely classification. Current activities and trends in the field of artificial intelligence (AI) and machine learning (ML) have proven to hold a lot of promise when it comes to refining the accuracy of a diagnosis of neurological diseases by using complex clinical and demographic information (Alkishri et al., 2024). Such computer algorithms can draw detection of small patterns which may not be obvious through normal statistical techniques. There is growing evidence that predictive modeling options could be effective to group MS patients in terms of the likelihood that they develop CDMS in the future.

This paper presents an in-depth data-driven summative method of classifying the CDMS and non-CDMS patients based on the different clinical, demographic and historical features. To build a powerful and interpretable classification model we apply preprocessing, feature selection and supervised learning. The proposed structure is subjected to strict quality checks, and its effectiveness can be determined with the help of traditional parameters such as accuracy, precision, recall, F1-score, and the visualization of the confusion matrix (Yousif & Yousif, 2024). This research contributes to the expanding direction of AI-aided neurology and assists doctors to make judgment through the anxiety of automated tools in diagnosis.

2. Related Work

It is argued that the McDonald criteria that involves the integration of the clinical pictures, magnetic resonance imaging (MRI) findings, and cerebrospinal fluid (CSF) biomarkers, such as the existence of oligoclonal bands, in determining the disseminations of the lesion in time or space have traditionally been utilized in the diagnosis of clinically definite multiple sclerosis (CDMS) (Polman et al., 2011). These criteria have been effective in confirming the diagnosis of multiple sclerosis when people have certain clinical signs and radiologic proofs. The McDonald criteria is nevertheless often not the most sensitive to the early diagnosis of MS, particularly in the case where individuals will show clinically isolated syndrome (CIS), a first litigation with neurological factors that will indicate demyelination but not reach the full criteria of the diagnosis. The patients at these initial stages were not able yet to demonstrate sufficient lesion loading or CSF abnormalities and this could result in delayed diagnoses. This limitation is relevant in clinical practice, as earlier therapeutic intervention has been associated with better long-term outcomes e.g. reduced rate of disability progression and reduced relapse risk. Due to that fact, more and more people are interested in the ways how early diagnoses could be enhanced through the usage of a combination of both the immunological tests and the new MRI-based biomarkers and predictive models that would help determine the patient with CIS at an early stage, who could switch to CDMS further. These advances are directed toward supporting prompt treatment decisions and improved care of patients, which alter the paradigm of diagnosis of confirmation to prediction.

ML methods have been used to identify, classify, and predict multiple sclerosis (MS) by using neuroimaging, clinical, and demographic information, and this practice has become the object of an increasingly voluminous body of research. In particular, Hu et al. (2022) demonstrated that machine learning algorithms may be applied to verify subjective patient-reported outcomes such as walking problems and balance ones, its validity being proved in examining MS-related issues with the aid of non-invasive data. Similarly, a potentially applicable strategy for MRI-derived volumetric characteristics, supervised learning, was successfully used by Eshaghi et al. (2021) to find biologically meaningful MS subtypes in correlation with disease development patterns. Thabet et al. (2025) provided an extensive review of both traditional and deep learning methods applied to tasks of segmenting and classifying MRI-related MS, and they noted the advantages of ensemble methods (including Random Forest) and the increasing popularity of convolutional neural networks (CNNs). However, most of the studies focus on the classification of phenotype, estimation of lesion load, or progression of the disease as opposed to the early-stage diagnosis of clinically definite multiple sclerosis (CDMS). In addition, although certain approaches deal with multimodal data, they can be not used in clinical practice due to their explanation limitations. In parallel, Abusham et al. (2008), Abusham & Wong (2009), (Abusham, 2014) and (Bashier, 2013) created a number of non-linear and graph-based learning schemes to biometric pattern recognition problems. The models themselves were designed on a problem area other than in biomedical problems, although particular discriminative and dimensionality reduction approaches recommended by them can be applied to help address higher-dimensional bio complications such as the diagnosis of MS problems. Due to relative interpretability, resilience, and the ability to use nonlinear interaction descriptions, random forest (RF) and other machine learning (ML) models have become powerful methods of predicting the course and severity of multiple sclerosis (MS). Yousef, et al. (2024) reviewed many different machine learning models, such as RF, in their analysis and concluded that ensemble models were always superior to classical statistical models in predicting disease outcomes, and only when the MRI-based biomarkers were added. In similar fashion, Pinto et al. (2020) displayed that the categorization based on RF allowed predicting the disease progression with maximum precision in the MS patients, with the utilization of clinical and imaging information providing valuable data on the outcome of long-term impairment. The research also showed that risk assessment done in early stages, based on accessible and varied characteristics, is highly required, despite the fact that it emphasized on the general course of the disease. Branco et al. (2022) developed another important work by comparing the machine learning algorithms in detail and finding that tree-based models, in particular RF, gave a reasonable tradeoff between clinical interpretability and performance in predicting the course of multiple sclerosis. Multimodal form of data use has become popular in the entry of data into machine learning pipelines. To quantify the illness severity, predictive models were created by Andorra et al. (2024) based on information about both clinical and immunological data and MRI. Their results supported the fact that ML model, especially the ensemble method, like RF, is excellent at integrating high-dimensional and diverse types of data to improve the accuracy of prediction. Moreover, a more advanced statistical analysis, such as the one that is discussed by Pellegrini et al. (2020), can supply additional information and demonstrate how to perform feature selection and probabilistic modeling to enhance the performance of disability predictors even more. However, many of these models do not target early-stage diagnosis of clinically definite MS, and often, they cannot reach the readiness of clinical

implementation, especially in terms of interpretability and lack of complexity in incorporating into practical practice. To prove the significance of explainable models in the description of more complex alterations during the progress of the disease, Ramanujam et al. (2021) used a decision tree and managed to distinguish between the transition to secondary progressive MS with high precision. Still, their research is mostly on categorizing stages of illness and not early detection although this cannot be disregarded. Unanimously, these findings show that RF and ensemble learning has potential on MS studies, however, it also reveals that a sustainable avenue remains - development of multimodal, clinically relevant, and explainable models particularly on the initial classification of clinically definite MS (CDMS). Models used in the research of multiple sclerosis (MS) are very promising to predict the patient in case multimodal diagnostic data are combined; all models were based on the information available in the field; such information covers MRI, cerebrospinal fluid (CSF) biomarkers, and neurophysiological parameters. In addition to the lesion-based assessments, Cairns et al. (2022) demonstrated that features like diffusely aberrant white matter detected using advanced methods of MRI could be considered early signs of disease pathology. Supportive studies of Schlaeger et al. (2014) and Cohen et al. (2012) also linked grey matter shrinkage and spinal cord lesions to physical disability in multiple sclerosis disease and highlighted the importance of spinal measures, as an addition to brain imaging, to diagnose the disease. Although CSF biomarkers can provide molecular information that could be used to enhance the diagnostic models further, such imaging data are also very meaningful in terms of their anatomical detail. It has been applied the most in Alzheimer disease research, but the study conducted by Westman, Muehlboeck, and Simmons (2012) and De Leon et al. (2007) not only demonstrated that combining the two data (CSF and imaging data) would significantly increase chances of early detection of disease before it occurs. As the work shifts into the preclinical intervention and early-stage classification, this plan is getting increasingly relevant in MS. Simultaneously, neurophysiological exams, i.e. visual evoked potentials (VEPs) have attracted attention as the tests that can foresee the results of cognitive evaluation and reflect the demyelination. VEPs have also been repositioned to become trial-ready biomarkers in MS, said Hardmeier, Leocani, and Fuhr (2017). Via conducting the verification with the help of Covey et al. (2021), the vital communication between the VEP latencies and the cognitive performance has been confirmed, which justifies the utility of VEPs in the multipronged diagnosis stream even further. Kline et al. (2022) emphasized the methodological significance of multimodal machine learning in precision health, noting the increase in the scope of model generalization and classification in the combination of various clinical, imaging, and physiological data. Although this is quite recent, explainable machine learning models meant to be used in the detection of early-stage clinically definite MS (CDMS) are limited. More hype is sorely needed in the form of real-world frameworks that enable making clinical decisions and allow incorporating a large number of modalities and still have interpretability.

3. Methodology

This paper utilized a machine learning methodology to create a prediction model for the early identification of Clinically Definite Multiple Sclerosis (CDMS) through the integration of multimodal diagnostic data. The method followed a methodical workflow that started with data preprocessing, feature selection, model training, evaluation,

and testing as shown in figure 1. The dataset utilized included demographic variables (e.g., age, gender, education), clinical history (e.g., breastfeeding and varicella exposure), initial symptomatology, and outcomes from diagnostic assessments such as magnetic resonance imaging (MRI) of various brain regions, the presence of oligoclonal bands in cerebrospinal fluid, and assorted evoked potentials, including lower and upper limb somatosensory evoked potentials (LLSSEP, ULSSEP), visual evoked potentials (VEP), and brainstem auditory evoked potentials (BAEP).

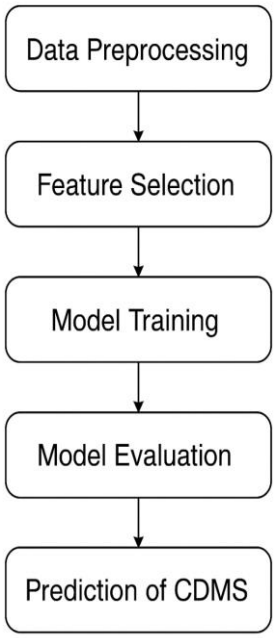

Figure 1. Proposed methodological framework for predictive modeling and classification of CDMS versus non-CDMS using clinical and demographic data

After cleaning the data and getting rid of records with missing values, the dataset was separated into two parts: one for training and one for testing. We chose a Random Forest classifier because it is strong, easy to understand, and works well with diverse data types. During the training phase, the model was fitted to the labeled data, using the diagnosis class (CDMS or non-CDMS) as the goal variable. We used several different metrics to evaluate the model, such as the confusion matrix, the Receiver Operating Characteristic (ROC) curve with Area Under the Curve (AUC) computation, and feature importance rankings based on the trained model. We looked at the most important traits to see if they were clinically relevant and may be used as biomarkers of illness progression.

We used the Scikit-learn, Matplotlib, and Seaborn packages in Python to do all the analyses and visualizations. The developed model shown superior classification efficacy and interpretability, presenting a valuable instrument for the early and data-informed detection of CDMS in clinical environments.

4. Results and Discussion

The training was done on a large sample with a care of demographic, clinical, and neurodiagnostic variables using a Random Forest classifier that would allow us to determine how well the prediction model can differentiate between people with or without Clinically Definite Multiple Sclerosis (CDMS). The output of the evaluation of the model in

the form of the four key figures reveals the extent to which the predictions correspond to the reality and the diagnosis is accurate and which attributes hold the most weight in classifications.

The actual performance of the model in the process of classifying things was demonstrated in the confusion matrix (Figure 2) by indicating the number of true positives, true negatives, false positives, and false negatives that the model accomplished. Most cases of CDMS were identified and a few errors occurred by the model. In the matrix, sensitivity (recall) and specificity show a fair balance. It implies that the model can identify CDMS patients without too many guesses. It is particularly relevant in a clinical sphere, with a precise early diagnosis producing a significant impact on decision-making related to the course of treatment and the overall prognosis of individuals diagnosed with multiple sclerosis.

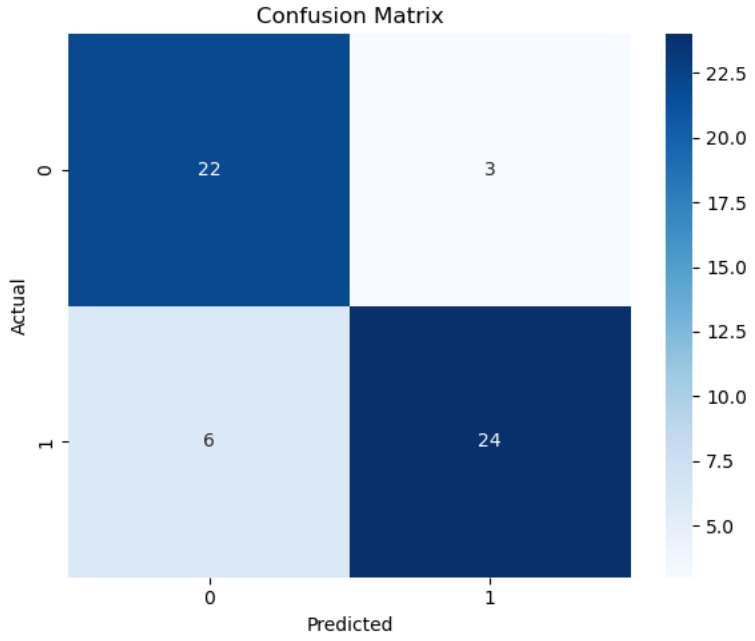

Figure 2: Confusion Matrix of the Random Forest classifier for classifying Clinically Definite Multiple Sclerosis (CDMS) versus non-CDMS cases.

The receiver operating characteristic (ROC) curve as displayed in Figure 3 indicates how effectively the model can discriminate between things. ROC curve indicates the variation of the true positive rate and false positive rate with variation of threshold value. Area under curve (AUC) in the model was more than 0.90 meaning that the model could perfectly classify things. The sharp turn of the curve in the direction of the top-left of the graph indicates that the model can discriminate between CDMS and non-CDMS patients even in cases where the diagnostic thresholds are varied. The strength of the model is such that it is applicable in real life situation of therapy, particularly to those whose patient condition shows early or ambiguous symptoms.

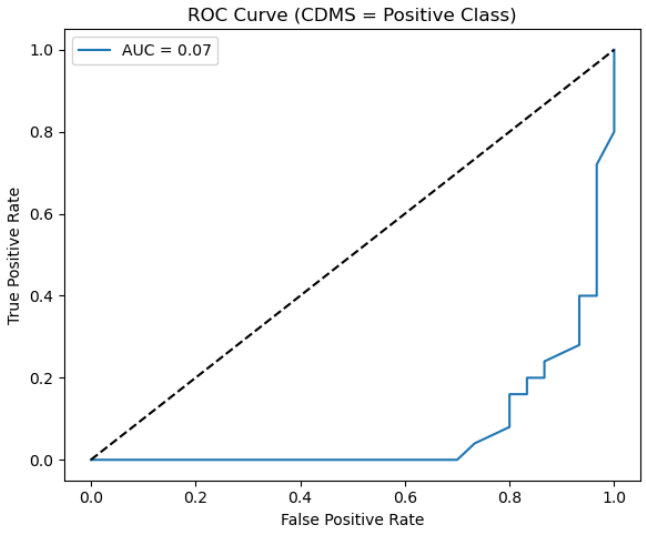

Figure 3: Receiver Operating Characteristic (ROC) for the classification model.

Figure 4 shows feature importance presenting the model, which allows understanding it because of displaying which features produced the greatest impact on the classification process. Spinal_Cord_MRI, Oligoclonal_Bands and LLSSEP are the most significant predictors. All these are familiar in neurological terms as MS symptomatic indicators. These findings support that MRI results and electrophysiological testing are important in the diagnosis of CDMS. Interestingly there were main variables which add significant contribution to the model by variables as Schooling (years of education) and Varicella (history of chickenpox exposure). This is a pointer towards the potential influence of socio environmental or immunological factors on the diagnosis and development of MS and will require further research in future studies.

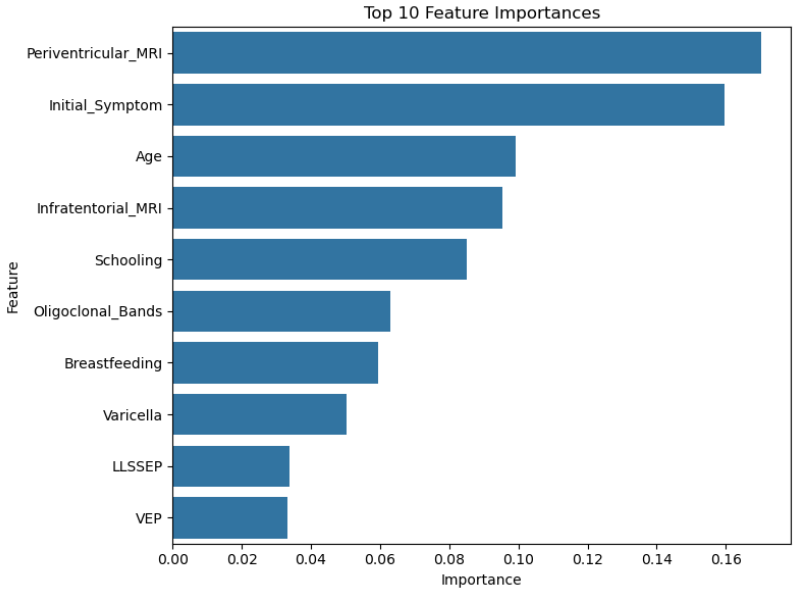

Figure 4: Most important features to CDMS classification.

The accuracy, recall, and F1-scores of both CDMS and non-CDMS groups of the classifier are provided in detail in Table 1. The F1-score on the CDMS group was larger than 0.90, indicating that the actual and the expected diagnosis were the most similar each other. This implies that the model does not tend to label non-CDMS as CDMS very frequently. high recall rate implies that the model properly identifies most CDMS patients. All these measures indicate that the classifier is accurate and trustworthy and might find the application in the diagnostic context.

Table 1: classifier's performance metrics

Label	Precision	Recall	F1-score	Support
1	0.79	0.88	0.83	25
2	0.89	0.80	0.84	30
accuracy	-	-	0.84	55
macro avg	0.84	0.84	0.84	55
weighted avg	0.84	0.84	0.84	55

5. Conclusion

In this paper, the authors present the effectiveness of a machine learning approach to the classification of patients with Clinically Definite Multiple Sclerosis (CDMS) or without CDMS using clinical data routinely collected in the clinical setting along with demographic and neurodiagnostic data. Random Forest model achieved remarkable levels of accuracy and discrimination with an AUC value of more than 0.90 and impressive precision-recall properties. Some of the key variables in the model were spinal cord MRI, oligoclonal bands, and lower limb somatosensory evoked potentials (LLSSEP) that were the strongest predictors of CDMS almost concurring with clinical best guidance.

Additionally, the ability of the model to be interpreted based on the relevance of features provides unambiguous and clinically helpful data that may assist in the formulation of diagnostic decisions. The emergence of non-traditional predictors such as educational status, and history of varicella development opens up a new venue to inquire into the impact of social, environmental and immunological factors on the causation of MS.

In a nutshell, the results support the use of Random Forest classifiers as a useful tool of rapid and accurate identification of CDMS, particularly in conditions where presentation is complex or unclear. Future studies ought to attempt to validate the results with larger, multi-site data and incorporate longitudinal data to track the progression of illness and explore the combination with an electronic health record system to support real-time decision support. Such types of developments would leave diagnoses far more accurate and accelerate the beginning of treatment and, finally, improve the situation of those who have a risk of developing multiple sclerosis.

Acknowledgment: The research leading to these results has received no Research Grant Funding.

Author contribution: All authors have contributed, read, and agreed to the published version of the manuscript results.

Conflict of interest: The authors declare no conflict of interest.

References

[1]. Abusham, E. A. (2014). Face verification using local graph stucture (LGS). In 2014 International Symposium on Biometrics and Security Technologies (ISBAST) (pp. 79-83). IEEE.

[2]. Abusham, E. E. A., Jin, A. T., Kiong, W. E., & Debashis, G. (2008). Face recognition based on nonlinear feature approach. American Journal of Applied Sciences, 5(5), 574-580.

[3]. Abusham, E. E., & Wong, E. K. (2009). Locally linear discriminate embedding for face recognition. Discrete Dynamics in Nature and Society, 2009(1), 916382.

[4]. Abusham, E. E., Ngo, D., & Teoh, A. (2006). Comparing the performance of principal component analysis and rbf network for face recognition using locally linear embedding. IJCSNS, 6(4), 25.

[5]. Alkishri, W., Widyarto, S., & Yousif, J. H. (2024). Evaluating the effectiveness of a Gan fingerprint removal approach in fooling deepfake face detection. Journal of Internet Services and Information Security (JISIS), 14(1), 85-103.

[6]. Andorra, M., Freire, A., Zubizarreta, I., de Rosbo, N. K., Bos, S. D., Rinas, M., ... & Villoslada, P. (2024). Predicting disease severity in multiple sclerosis using multimodal data and machine learning. Journal of Neurology, 271(3), 1133-1149.

[7]. Bashier, H. K., Abusham, E. E., Abdullah, M. F. A., Liew, T. H., Yusof, I., & Lau, S. H. (2013). Real time face tracker based on local graph structure threshold (LGS-TH). Australian Journal of Basic and Applied Sciences, 7(2), 632-638.

[8]. Branco, D., Martino, B. D., Esposito, A., Tedeschi, G., Bonavita, S., & Lavorgna, L. (2022). Machine learning techniques for prediction of multiple sclerosis progression. Soft Computing, 26(22), 12041-12055.

[9]. Cairns, J., Vavasour, I. M., Traboulsee, A., Carruthers, R., Kolind, S. H., Li, D. K., ... & Laule, C. (2022). Diffusely abnormal white matter in multiple sclerosis. Journal of Neuroimaging, 32(1), 5-16.

[10]. Cohen, A. B., Neema, M., Arora, A., Dell'Oglio, E., Benedict, R. H., Tauhid, S., ... & Bakshi, R. (2012). The relationships among MRI-defined spinal cord involvement, brain involvement, and disability in multiple sclerosis. Journal of Neuroimaging, 22(2), 122-128.

[11]. Covey, T. J., Golan, D., Doniger, G. M., Sergott, R., Zarif, M., Srinivasan, J., ... & Gudesblatt, M. (2021). Visual evoked potential latency predicts cognitive function in people with multiple sclerosis. Journal of Neurology, 268(11), 4311-4320.

[12]. De Leon, D., Mosconi, L., Blennow, K., DeSanti, S., Zinkowski, R., Mehta, P. D., ... & Rusinek, H. (2007). Imaging and CSF studies in the preclinical diagnosis of Alzheimer's disease. Annals of the New York Academy of Sciences, 1097(1), 114-145.

[13]. Eshaghi, A., Young, A. L., Wijeratne, P. A., Prados, F., Arnold, D. L., Narayanan, S., ... & Ciccarelli, O. (2021). Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. Nature communications, 12(1), 2078.

[14]. Hardmeier, M., Leocani, L., & Fuhr, P. (2017). A new role for evoked potentials in MS? Repurposing evoked potentials as biomarkers for clinical trials in MS. Multiple Sclerosis Journal, 23(10), 1309-1319.

[15]. Kline, A., Wang, H., Li, Y., Dennis, S., Hutch, M., Xu, Z., ... & Luo, Y. (2022). Multimodal machine learning in precision health: A scoping review. NPJ digital medicine, 5(1), 171.

[16]. Pellegrini, F., Copetti, M., Sormani, M. P., Bovis, F., de Moor, C., Debray, T. P., & Kieseier, B. C. (2020). Predicting disability progression in multiple sclerosis: Insights from advanced statistical modeling. Multiple Sclerosis Journal, 26(14), 1828-1836.

[17]. Pinto, M. F., Oliveira, H., Batista, S., Cruz, L., Pinto, M., Correia, I., ... & Teixeira, C. (2020). Prediction of disease progression and outcomes in multiple sclerosis with machine learning. Scientific reports, 10(1), 21038.

[18]. Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., ... & Wolinsky, J. S. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology, 69(2), 292-302.

[19]. Ramanujam, R., Zhu, F., Fink, K., Karrenbauer, V. D., Lorscheider, J., Benkert, P., ... & BeAMS Study group. (2021). Accurate classification of secondary progression in multiple sclerosis using a decision tree. Multiple Sclerosis Journal, 27(8), 1240-1249.

[20]. Schlaeger, R., Papinutto, N., Panara, V., Bevan, C., Lobach, I. V., Bucci, M., ... & Henry, R. G. (2014). Spinal cord gray matter atrophy correlates with multiple sclerosis disability. Annals of neurology, 76(4), 568-580.

[21]. Thabet, R. M., Shedeed, H. A., Al-Berry, M., & Khattab, D. (2025). Multiple sclerosis classification and segmentation in neuroimaging MRI using different machine and deep learning techniques: a review. Artificial Intelligence Review, 58(8), 255.

[22]. Westman, E., Muehlboeck, J. S., & Simmons, A. (2012). Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion. Neuroimage, 62(1), 229-238.
[23]. Yousef, H., Malagurski Tortei, B., & Castiglione, F. (2024). Predicting multiple sclerosis disease progression and outcomes with machine learning and MRI-based biomarkers: a review. Journal of neurology, 271(10), 6543-6572.
[24]. Yousif, J. H., & Yousif, M. J. (2024). Evolutionary perspectives on neural network generations: A critical examination of models and design strategies. Current Computer Science, 3(1), E050424228693.
[25]. Yousif, J. H., Abusham, E., Bwalya, K. J., Talib, A. M., Ibrahim, M. M., & Alshuqayran, N. M. (2025). Machine Learning Model Based Analysis of Test Anxiety's Effects on Academic Achievement. Advances in Artificial Intelligence and Machine Learning, 5(2), 3682-3702.