American Journal of Sciences and Engineering Research

E-ISSN -2348 – 703X, Volume 7, Issue 5, 2024

Machine Learning in Diagnosis of Multiple Sclerosis: Identification of Risk Factors and Performance Evaluation with XGBoost

Ipek Balikci Cicek[1], Zeynep Kucukakcali[2]

[1]*Inonu University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Malatya, Turkey.*
[1,2] *Department of Biostatistics and Medical Informatics, Inonu University Faculty of Medicine, 44280, Malatya, Turkey.*

ABSTRACT:

Aim: *Multiple sclerosis (MS) is a chronic autoimmune disease characterized by the degradation of myelin in the central nervous system, leading to various neurological symptoms that can significantly impact the quality of life. The complexity of MS pathogenesis stems from an interplay of genetic, environmental, and immunological factors, resulting in inflammation and neurodegeneration. Due to its economic burden and the potential severity of its progression, early and accurate prediction of MS has become increasingly vital. Therefore, this study is by leveraging an open-access dataset comprising patient data, both with and without MS, the study aims to identify potential risk factors and improve the understanding of the disease.*

Material and Method: *This study employed Extreme Gradient Boosting (XGBoost), a robust and adaptable machine learning (ML) algorithm, to predict the occurrence of MS. A publicly accessible dataset comprising 100 patients (50 with MS and 50 healthy controls) was used to identify potential risk factors. The mean age of the participants was 45.68±11.71 years, with 74% female and 26% male. A 5-fold cross-validation method was applied to ensure the model's robustness and generalizability, evaluating various performance metrics, including accuracy, sensitivity, specificity, positive predictive value, and F1-score. All modeling and computations were conducted using Python.*

Results: *The XGBoost model demonstrated strong performance, achieving an accuracy of 94.0%, sensitivity of 92.0%, and specificity of 96.0%. The analysis revealed statistically significant differences between MS patients and controls in variables such as superior sagittal sinus (SSS) area and circumference, right transverse sinus area, left sigmoid sinus area and circumference, left distal sigmoid sinus area and circumference, and right and left optic nerve sheath diameters. Feature importance analysis indicated that the most significant variables for MS prediction were the right transverse sinus area, SSS circumference, SSS area, left transverse sinus area, and left optic nerve sheath diameter.*

Conclusion: *These findings enhance the understanding of MS, highlighting critical predictors that could contribute to early diagnosis and more effective risk management strategies, underscoring the transformative potential of ML in modern medicine.*

Keywords: *Multiple Sclerosis, Machine Learning, Extreme Gradient Boosting, Classification Modeling, Disease Prediction, Risk Factors*

I. INTRODUCTION

Multiple sclerosis (MS) is a chronic autoimmune illness marked by the degradation of myelin, the protective coating that surrounds nerve fibers in the central nervous system (CNS). This demyelination causes a variety of neurological symptoms, including motor and cognitive deficits, sensory problems, and weariness, all of which can have a major impact on the quality of life of those afflicted (1, 2). The illness is divided into numerous clinical variants, the most common of which are relapsing-remitting multiple sclerosis, secondary progressive multiple sclerosis, and primary progressive multiple sclerosis, each with its own set of progression patterns and symptoms. MS pathogenesis is characterized by intricate interplay between genetic predispositions, environmental variables, and immune system dysregulation, which lead to inflammation and neurodegeneration (3, 4). Understanding multiple sclerosis is critical not just for therapeutic reasons, but also because it offers substantial difficulties to healthcare systems and society at large. MS is connected with significant economic consequences due to healthcare expenditures, decreased productivity, and the need for long-term care (5).

Moreover, the typical progression of MS might vary from moderate to significantly severe. The majority of individuals experience an initial relapsing-remitting phase characterized by the intermittent manifestation of symptoms. Ultimately, remissions cease, and the condition progresses to a secondary progressive form, resulting in heightened impairment (6). This is a significant concern as the condition can be profoundly debilitating, and several existing treatments may have side effects that are more severe than the illness itself. Consequently, the ability to generate highly accurate predictions is becoming increasingly valuable (7).

Machine learning (ML) has emerged as a transformative technology across various domains, particularly in healthcare, where it plays a pivotal role in diagnostics and treatment planning. The foundational principles of ML involve various tasks such as classification, regression, and clustering, which are essential for developing effective predictive models. provide a comprehensive overview of these tasks, indicating that the improvement of ML algorithms is contingent upon their ability to learn from experience and adapt to new data (8). This adaptability is crucial in medical applications where patient data can be highly variable. further elaborates on the capabilities of ML algorithms to detect patterns in data, underscoring their utility in diverse applications, including healthcare (9). The evolution of these algorithms has led to the development of sophisticated techniques that can process large datasets, thereby enhancing the accuracy of predictions. Moreover, the choice of algorithms plays a significant role in the performance of ML models. discusses the importance of feature coding and the selection of appropriate algorithms in the learning process, which is critical for achieving optimal results in various applications, including medical diagnostics(10). This is echoed by , who categorizes ML algorithms into supervised, unsupervised, and reinforcement learning, each serving distinct purposes in data analysis (11). In conclusion, the application of ML in various fields, especially healthcare, underscores its transformative potential. The ability of ML algorithms to learn from data, adapt to new information, and improve diagnostic accuracy positions them as invaluable tools in modern medicine. However, ongoing research is necessary to refine these algorithms, address their limitations, and ensure their effective integration into clinical practice.

Extreme Gradient Boosting (XGBoost) has developed as a robust and adaptable ML algorithm, well-known for its efficiency and efficacy in a variety of predictive modeling applications. XGBoost, created by Tianqi and Carlos Guestrin, is an improvement on existing gradient boosting algorithms that includes various modifications that increase speed and performance. The approach works by generating an ensemble of decision trees in a sequential order, with each new tree correcting the mistakes of the preceding ones, iteratively refining the model (12, 13). One of XGBoost's important strengths is its capacity to tolerate missing information and its resistance to overfitting, making it suited for a broad range of applications, including finance and healthcare (14, 15).

In this study, the aim was to predict the occurrence of multiple sclerosis (MS) and identify the associated risk factors by applying the XGBoost method to a dataset comprising open-access patient data, including both individuals with and without MS. Through this approach, the study seeks to enhance the understanding of MS by uncovering potential predictors and contributing to early diagnosis and effective risk management strategies.

II. MATERIAL AND METHOD

Dataset and Variables

The"Multiple Sclerosis" dataset to be used in this study is available as open access at https://data.mendeley.com/datasets/mytp5z2zdd. There are 100 patients in the dataset in total. While 50 of these patients are Multiple Sclerosis patients, the remaining 50 are in the control group. The variables in the dataset and their characteristics are presented in detail in Table 1.

Table 1: The variables included in the dataset

Variables	Variable Types	Variable Roles
Sex (Male/Female)	Qualitative	Predictor
Weight (kg)	Quantitative	Predictor
Height (m)	Quantitative	Predictor
Age at Visit (years)	Quantitative	Predictor
Disease Duration (Yrs)	Quantitative	Predictor
Expanded Disability Status Scale (EDSS)	Quantitative	Predictor
Superior Sagittal Sinüs (SSS) Area (mm^2)	Quantitative	Predictor
SSS Circumfrence (mm)	Quantitative	Predictor
Right Transverse Sinus Area (mm^2)	Quantitative	Predictor
Right Transverse Sinus Circumfrence (mm)	Quantitative	Predictor
Left Transverse Sinus Area (mm^2)	Quantitative	Predictor
Left Transverse Sinus Circumfrence (mm)	Quantitative	Predictor
Right Sigmoid Area (mm^2)	Quantitative	Predictor
Right Sigmoid Circumfrence (mm)	Quantitative	Predictor
Left Sigmoid Area (mm^2)	Quantitative	Predictor
Left Sigmoid Circumfrence (mm)	Quantitative	Predictor
Right Distal Sigmoid Area (mm^2)	Quantitative	Predictor
Right Distal Sigmoid Circumfrence (mm)	Quantitative	Predictor
Left Distal Sigmoid Area (mm^2)	Quantitative	Predictor
Left Distal Sigmoid Circumfrence (mm)	Quantitative	Predictor
Right Optic Nerve Sheath Diameter (mm)	Quantitative	Predictor
Left Optic Nerve Sheath Diameter (mm)	Quantitative	Predictor
Pituitary Height (mm)	Quantitative	Predictor
Group (Multiple Sclerosis/Control)	Qualitative	Output

Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is a powerful ML algorithm that has gained significant attention due to its high performance in various predictive modeling tasks. Developed as an enhancement of the traditional Gradient Boosting Decision Trees (GBDT), XGBoost incorporates several optimizations that improve its speed and accuracy, making it suitable for large-scale data applications (13). The algorithm's design allows it to efficiently handle missing values and complex data structures, which is particularly beneficial in fields such as healthcare, finance, and environmental science (16, 17). One of the key advantages of XGBoost is its ability to model non-linear relationships effectively. This capability has been demonstrated in various studies, including its application in predicting outcomes in critical care settings, such as acute kidney injury and mortality rates among ICU patients (18).

XGBoost's performance is attributed to its ensemble learning approach, where it combines the predictions of multiple weak learners (decision trees) to produce a stronger overall model. This method not only enhances predictive accuracy but also improves the model's generalization capabilities, making it less prone to overfitting (13, 19). The algorithm's scalability allows it to be applied to large datasets efficiently, which is crucial in today's

data-driven environments(19, 20). Additionally, XGBoost provides valuable insights into feature importance, enabling practitioners to understand the contributions of various predictors in their models (21). Despite its strengths, XGBoost is not without limitations. The algorithm can exhibit poor interpretability due to its complex structure, which may pose challenges in applications where understanding the decision-making process is critical (22). Nevertheless, ongoing research aims to enhance the interpretability of XGBoost models while maintaining their predictive power (23). Furthermore, the hyperparameter tuning process can be intricate, requiring careful optimization to achieve the best performance (20). In conclusion, XGBoost stands out as a leading ML algorithm due to its efficiency, accuracy, and versatility across various applications. Its ability to handle complex data relationships and provide insights into feature importance makes it a valuable tool for researchers and practitioners alike. As the field of ML continues to evolve, XGBoost is likely to remain a prominent choice for predictive modeling tasks.

III. Biostatistical Data Analyses

Qualitative variables included in the study are expressed using numbers and percentages. The Shapiro-Wilk test was used to evaluate whether the quantitative data followed a normal distribution. Data that did not show normal distribution were presented using median (minimum-maximum), while normal distribution was summarized by mean ± standard deviation. In statistical analysis, categorical variables were evaluated using the Continuity Correction Chi-square test. To examine quantitative variables between two independent groups, Mann Whitney U test and Independent samples t-test was used. In statistical analysis, a p value below 0.05 was considered statistically significant. All analyses were performed using IBM SPSS Statistics 26.0 for Windows (New York, USA).

ML Modeling and Performance Evaluation

The research utilized the XGBoost methodology during the modeling phase, implementing it on the entire dataset to develop a predictive model. XGBoost is a prevalent machine learning technique noted for its effectiveness, versatility, and exceptional performance in handling structured data. The n-fold cross-validation technique was employed for the analysis to ensure the model's resilience and generalizability. Cross-validation is a statistical method utilized to evaluate the performance of machine learning models. In n-fold cross-validation, the entire dataset is randomly divided into n equally sized subsets, known as "folds." The model is then trained n times, using a different fold as the testing set and the remaining n-1 folds as the training set for each iteration. This iterative method ensures that each data subset is employed for both training and validation, enabling a more comprehensive evaluation of the model's efficacy. This study utilized a 5-fold cross-validation technique, in which the dataset was divided into five pieces. During each iteration, four segments of the data were employed for model training, while the remaining segment was designated for testing. The method was conducted five times, with each segment employed as the test set once. The results from each iteration were averaged to obtain a more reliable evaluation of the model's performance metrics. Various metrics were employed to assess the model's performance, including accuracy, balanced accuracy, sensitivity, specificity, positive predictive value, negative predictive value, and the F1-score. These metrics provide a comprehensive evaluation of the model's performance across all aspects, including its ability to precisely identify positive instances (sensitivity) and its effectiveness in correctly detecting negative cases (specificity). The F1-score is the harmonic mean of precision and recall, offering a balanced evaluation of the model's accuracy in scenarios with imbalanced input. The research additionally analyzed the significance of each input variable within the model. Variable significance, or feature importance, indicates the extent to which each input variable affects the model's predictions. Recognizing the most significant aspects facilitates the understanding of the model's decision-making process and provides essential insights into the underlying data patterns. All modeling and computations were performed using the Python programming language (24).

IV. RESULTS

The data set used in this study included a total of 100 patients; 50 of these patients were Multiple Sclerosis patients and the remaining 50 were healthy individuals in the control group. The mean age of the patients in the study was 45.68±11.71 years. When analyzed in more detail, the mean age of Multiple Sclerosis patients was 46.44±12.51 years, while the mean age of patients in the control group was 44.91±10.93 years. In terms of gender distribution, 74 (74%) of the participants were female and 26 (26%) were male.

When the data on the disease duration of Multiple Sclerosis patients were analyzed, the median duration of the disease was found to be 10.79 years, ranging from 0.07 years to 28.93 years. This reveals that the duration of the disease shows a significant variability among individuals. In addition, the median score on the Expanded Disability Status Scale of Multiple Sclerosis patients was 2.75, ranging from 0 to 8.5.

The results of the statistical analyzes of the independent variables in terms of the target variable are given in Table 2.

Table 2: Statistical Analysis Results Between the Target Variable and Independent Variables

Variables		Group				*p*
		MS		Control		
Sex n (%)	Female	37 (74)		37 (74)		1.0***
	Male	13 (26)		13 (26)		
		Mean±SD	Median(Min-Max)	Mean±SD	Median(Min-Max)	
Weight (kg)		74.34±17.43	72.5(45-116)	79.4±21.54	74.5(53-140)	0.346**
Height (m)		1.67±0.08	1.66(1.5-1.83)	1.67±0.11	1.68(1.16-1.83)	0.523**
Age at Visit (years)		46.44±12.51	46.35(21.7-73.3)	44.91±10.93	47.75(20.9-62.2)	0.807**
SSS area (mm^2)		49.51±11.31	47.05(29.1-78.8)	40.97±11.19	38.5(23.4-75.5)	**<0.001****
SSS Circumfrence (mm)		32.36±5.32	32.25(23.8-44.4)	27.53±4.72	26.45(20.4-45)	**<0.001****
Right Transverse Sinus Area (mm^2)		29.11±15.43	25(4.2-69.8)	38.79±13.92	37.5(6.2-67.8)	**<0.001****
Right Transverse Sinus Circumfrence (mm)		26.19±7.67	25.2(9.9-51.6)	26.53±6.07	26.5(11.6-40)	0.432**
Left Transverse Sinus Area (mm^2)		25.59±14.08	22.55(5.9-66.2)	27.81±12.6	27(8.9-69.5)	0.285**
Left Transverse Sinus Circumfrence (mm)		25.16±9.75	23(10-56.4)	22.99±5.81	23.45(12-41.6)	0.457**
Right Sigmoid Area (mm^2)		48.58±17.56	46.1(21.4-113.7)	49.74±18.32	48.05(15.3-119.1)	0.533**
Right Sigmoid Circumfrence (mm)		30.77±5.82	29.75(20.5-55.3)	29.64±5.58	29.2(15.9-50.9)	0.296**
Left Sigmoid Area (mm^2)		46.39±19.03	45.05(4.8-116.9)	37.88±15.52	34.65(15-79.5)	**0.006***
Left Sigmoid Circumfrence (mm)		28.74±6.41	29.1(8.4-44.1)	24.86±4.85	24.2(15.5-37.4)	**0.001***

Right Distal Sigmoid Area (mm^2)	51.41±23.7	49.7(12-112.9)	53.77±21.04	53.35(21.4-109.8)	0.599*
Right Distal Sigmoid Circumfrence (mm)	27.45±6.35	27.55(12.5-43.5)	27.93±5.26	28.6(17.2-38.3)	0.685*
Left Distal Sigmoid Area (mm^2)	50.07±23.28	44.05(7.9-133.1)	37.79±16.72	37.4(5.2-80.9)	**0.007****
Left Distal Sigmoid Circumfrence (mm)	26.69±5.8	26.45(12.7-40.9)	23.18±5.81	23.5(9.6-39.8)	**0.003***
Right Optic Nerve Sheath Diameter (mm)	6.21±0.91	6.2(4.6-8.4)	5.52±0.69	5.7(4.1-6.8)	**<0.001***
Left Optic Nerve Sheath Diameter (mm)	5.99±0.89	5.9(4.1-7.8)	5.35±0.67	5.2(4.2-6.8)	**<0.001****
Pituitary Height (mm)	5.09±1.37	5(2.4-8.1)	5.09±1.37	5.05(2.4-8.7)	0.983*

***: Independent samples t-test, **: Mann Whitney U test, ***: Continuity Correction test, Min: Minimum, Max: Maximum, SD: Standart Deviation**

Based on the analysis of the relationship between the independent variables and the target variable presented in Table 2, there is a statistically significant difference ($p < 0.05$) between the multiple sclerosis patients and the control group in several measured parameters. Specifically, these parameters include the SSS area (mm^2) and SSS circumference (mm), right transverse sinus area (mm^2), left sigmoid sinus area (mm^2) and its circumference (mm), left distal sigmoid sinus area (mm^2) and its circumference (mm), as well as the right and left optic nerve sheath diameters (mm). This suggests that these variables show notable variations between the two groups, indicating potential associations with the presence of multiple sclerosis.

Table 3: Performance Metrics of the XGBoost Model

Performance Metrics	Value (%)
Accuracy (Acc)	94.0
Balanced Accuracy	94.0
Sensitivity	92.0
Specificity	96.0
Positive predictive value	95.8
Negative predictive value	92.3
F1-score	93.9

This table summarizes the performance of the XGBoost model, providing a comprehensive overview of its accuracy, sensitivity, specificity, and predictive values, indicating a strong performance in distinguishing between the different classes. Specifically, the model achieved an accuracy of 94.0% and a balanced accuracy of 94.0%, reflecting its overall performance in correctly classifying the data. The sensitivity was 92.0%, highlighting the model's effectiveness in identifying true positive cases. The specificity was 96.0%, indicating a high rate of correctly identifying true negatives. Furthermore, the positive predictive value was 95.8%, while the negative predictive value stood at 92.3%, demonstrating the model's reliability in predicting both positive and negative outcomes. Finally, the F1-score reached 93.9%, offering a harmonic mean of the model's precision and sensitivity, thereby reflecting its overall accuracy in handling both false positives and false negatives.

Figure 1 presents a graphic displaying the performance metric values obtained from XGBoost model. These performance metrics have been evaluated based on various criteria to analyze the accuracy and success levels of the models.

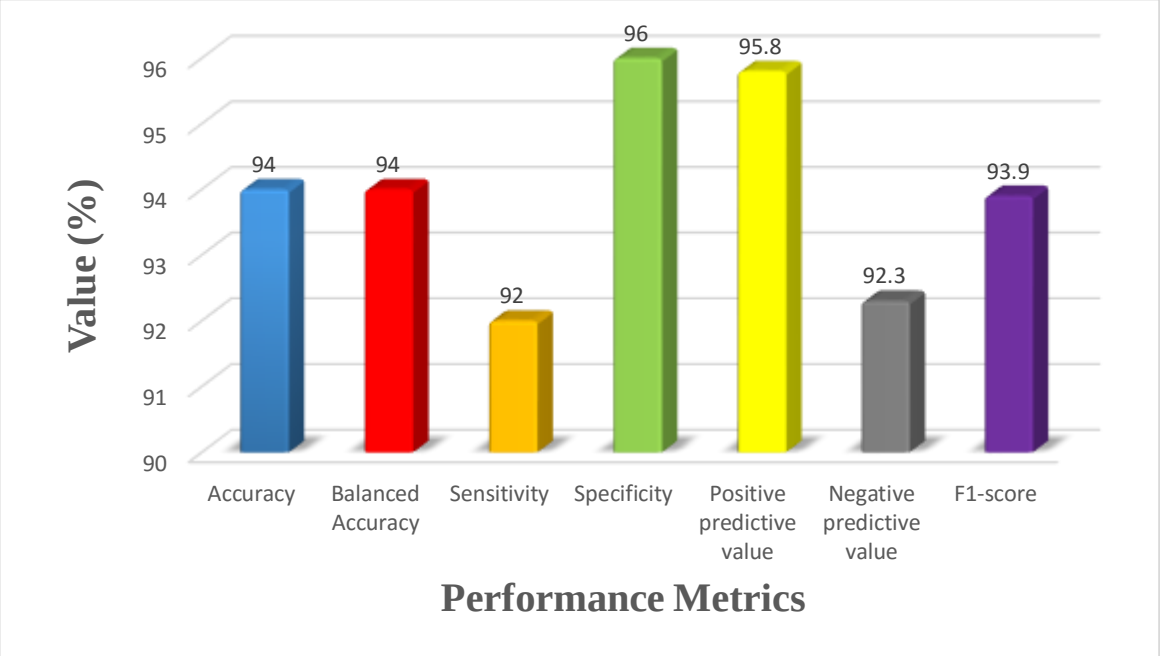

Figure 1: The performance metrics obtained from the XGBoost model."

Table 4 provides an overview of the variable importance values derived from the XGBoost model. This table highlights the relative significance of each variable in the model, illustrating how different features contribute to the model's predictions and overall performance."

Table 4: Importance values for variables determined by the XGBoost model

Variables	Importance Values
Right Transverse Sinus Area (mm^2)	100
Sss Circumfrence (mm)	92.958
Sss Area Mm2	78.444
Left Transverse Sinus Area (mm^2)	59.856
Left Optic Nerve Sheath Diameter (mm)	47.273
Left Distal Sigmoid Area (mm^2)	39.159
Left Transverse Sinus Circumfrence (mm)	34.618
Right Optic Nerve Sheath Diameter (mm)	28.144
Left Sigmoid Circumfrence (mm)	24.181
Right Sigmoid Circumfrence (mm)	11.759
Age At Visit (Years)	11.24

Figure 2 presents a graph illustrating the importance values of variables as determined by the XGBoost model. The analysis revealed that the top five most significant variables associated with multiple sclerosis are: the area of the right transverse sinus (in mm^2), the circumference of the superior sagittal sinus (SSS) (in mm), the area of the SSS (in mm^2), the area of the left transverse sinus (in mm^2), and the diameter of the left optic nerve sheath (in mm).

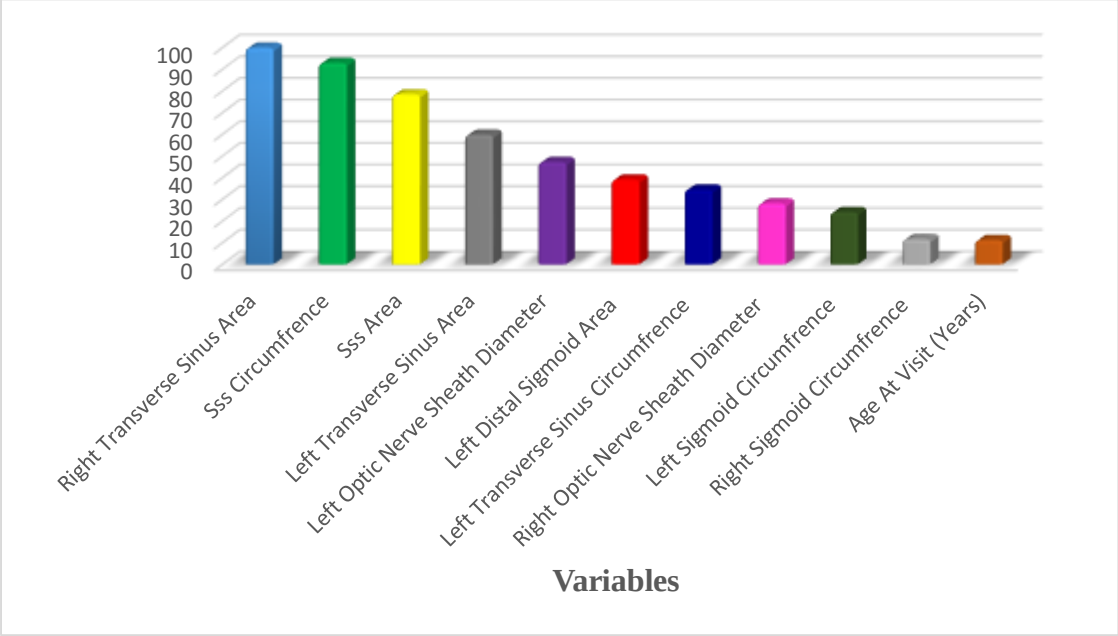

Figure 2: The importance values of the variables obtained as a result of the XGBoost model

V. DISCUSSION

MS is a neurodegenerative illness that primarily affects the central nervous system, specifically the brain, spinal cord, and optic nerve (25, 26). MS is difficult to diagnose due to the fact that the most prevalent diagnostic strategy is mostly centered on eliminating other diseases utilizing paraclinical tests that are frequently time-consuming, expensive, and intrusive. Furthermore, treatments are administered without regard for each MS patient's disability course (27). In light of these challenges, this study aims to predict the likelihood of MS and identify associated risk factors using the XGBoost ML method on an open-access dataset comprising both MS patients and healthy controls.

As a result of the modeling process, the performance metrics obtained were as follows: accuracy was 94.0%, balanced accuracy was 94.0%, sensitivity was 92.0%, specificity was 96.0%, positive predictive value was 95.8%, negative predictive value was 92.3%, and the F1-score was 93.9%.

Based on the variable importance values derived from the model, the five most significant variables associated with multiple sclerosis were identified as follows: the area of the right transverse sinus (measured in square millimeters), the circumference of the superior sagittal sinus (measured in millimeters), the area of the superior sagittal sinus (measured in square millimeters), the area of the left transverse sinus (measured in square millimeters), and the diameter of the left optic nerve sheath (measured in millimeters).

Additionally, in the context of MS, our study has identified statistically significant differences in various parameters between MS patients and a control group, with a p-value of less than 0.05. The parameters measured include the SSS area and circumference, right transverse sinus area, left sigmoid sinus area and circumference, left distal sigmoid sinus area and circumference, as well as the right and left optic nerve sheath diameters. These findings are consistent with existing literature that highlights the neuroanatomical and physiological alterations associated with MS, suggesting that these variables may serve as potential biomarkers for the disease.

The SSS area and circumference, along with the transverse sinus measurements, have been implicated in the pathophysiology of MS. For instance, studies have shown that venous abnormalities, including those in the transverse and sigmoid sinuses, can be prevalent in MS patients, potentially contributing to the disease's progression and symptomatology (28, 29). Furthermore, the optic nerve sheath diameter is a critical parameter in assessing optic nerve involvement in MS, as increased diameters have been associated with optic neuritis, a common manifestation of the disease (30, 31). This correlation underscores the importance of these measurements in understanding the structural changes that occur in MS.

Moreover, the variations observed in the left sigmoid sinus and its distal segment may reflect underlying venous drainage issues that have been hypothesized to exacerbate MS symptoms. Research indicates that impaired venous drainage can lead to increased intracranial pressure, which may further complicate the clinical picture of MS (32, 33). The significance of these findings lies not only in their statistical relevance but also in their potential implications for clinical practice, where such measurements could aid in the diagnosis and monitoring of MS progression.

In addition to the anatomical parameters, the relationship between MS and various psychological and cognitive factors has also been documented. For example, studies have shown that MS patients often experience higher levels of social anxiety and depression, which can significantly impact their quality of life (34, 35). The interplay between neuroanatomical changes and psychological well-being is crucial, as cognitive impairments can further complicate the management of MS, necessitating a comprehensive approach to treatment that addresses both physical and mental health (36, 37).

The findings from our study align with previous research that has established a link between structural brain changes and cognitive dysfunction in MS patients. For instance, neuroimaging studies have demonstrated that lesions in specific brain regions correlate with cognitive deficits, emphasizing the need for early detection and intervention strategies (38, 39). This highlights the importance of integrating neuroimaging findings with clinical assessments to provide a holistic view of the patient's health status.

Additionally, the role of environmental and genetic factors in the development of MS cannot be overlooked. Epidemiological studies have indicated that lifestyle factors, such as obesity and physical inactivity, may increase the risk of developing MS, particularly in young adults (38, 40). This suggests that preventive strategies targeting modifiable risk factors could play a significant role in reducing the incidence of MS and improving patient outcomes.

Furthermore, the association between MS and autoimmune processes has been extensively studied. The dysregulation of immune responses, particularly involving T-cells and B-cells, has been implicated in the pathogenesis of MS (41, 42). Understanding these immunological mechanisms is essential for developing targeted therapies that can modify the disease course and alleviate symptoms.

In conclusion, the statistically significant differences observed in our study between MS patients and the control group across various measured parameters provide valuable insights into the disease's pathophysiology. The implications of these findings extend beyond mere statistical significance, as they highlight the need for a multifaceted approach to MS management that encompasses neuroanatomical, psychological, and immunological factors. Future research should continue to explore these relationships to enhance our understanding of MS and improve patient care.

VI. REFERENCES

1. Caravagna C. What Is Multiple Sclerosis? Frontiers for Young Minds. 2019;7.
2. Olkiewicz J, Bonek R, Filipska K, Ślusarz R. Adherence to Therapeutic Recommendations in Patients Suffering From Multiple Sclerosis. The Journal of Neurological and Neurosurgical Nursing. 2020;9(3):103-7.
3. Dymecka J, Bidzan M. Biomedical Variables and Adaptation to Disease and Health-Related Quality of Life in Polish Patients With MS. International Journal of Environmental Research and Public Health. 2018;15(12):2678.
4. Cocchi M, Mondo E, Romeo M, Traina G. The Inflammatory Conspiracy in Multiple Sclerosis: A Crossroads of Clues and Insights Through Mast Cells, Platelets, Inflammation, Gut Microbiota, Mood Disorders and Stem Cells. International Journal of Molecular Sciences. 2022;23(6):3253.
5. Havrdová E, Kobelt G, Berg J, Capsa D, Gannedahl M, Doležal T. New Insights Into the Burden and Costs of Multiple Sclerosis in Europe: Results of the Czech Republic. Multiple Sclerosis Journal. 2017;23(2_suppl):41-52.
6. Wottschel V, Alexander D, Kwok P, Chard D, Stromillo M, De Stefano N, et al. Predicting outcome in clinically isolated syndrome using machine learning. NeuroImage: Clinical. 2015;7:281-7.
7. Tacchella A, Romano S, Ferraldeschi M, Salvetti M, Zaccaria A, Crisanti A, et al. Collaboration between a human group and artificial intelligence can improve prediction of multiple sclerosis course: a proof-of-principle study. F1000Research. 2017;6.

8. Htun TM, Tun Z. Algorithm Tuning From Comparative Analysis of Classification Algorithms. International Journal of Research and Studies Publishing. 2018;8(5).
9. Li Z, Zhang F, Wang Y. Research on the Application of Machine Learning in Smart Finance. 2023.
10. Wang J. Fraud Detection in Digital Payment Technologies Using Machine Learning. Journal of Economic Theory and Business Management. 2024;1(2):1-6.
11. Makkar A, Santosh K. SecureFed: federated learning empowered medical imaging technique to analyze lung abnormalities in chest X-rays. International Journal of Machine Learning and Cybernetics. 2023;14(8):2659-70.
12. Kurugama KM, Kazama S, Hiraga Y, Samarasuriya C. A Comparative Spatial Analysis of Flood Susceptibility Mapping Using Boosting Machine Learning Algorithms in Rathnapura, Sri Lanka. Journal of Flood Risk Management. 2024;17(2).
13. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. 2016.
14. Quinino RMe, Agena F, Andrade LGMd, Borba MF, Filho ADPC, David-Neto E. A Machine Learning Prediction Model for Immediate Graft Function After Deceased Donor Kidney Transplantation. Transplantation. 2023;107(6):1380-9.
15. Ahamad MM, Aktar S, Uddin MJ, Rahman T, Alyami SA, Al-Ashhab S, et al. Early-Stage Detection of Ovarian Cancer Based on Clinical Data Using Machine Learning Approaches. Journal of Personalized Medicine. 2022;12(8):1211.
16. Hakam A, Utama W, Garini SA, Jabar OA, Insani ANF, Rosandi Y. Sonic Log Prediction Based on Extreme Gradient Boosting (XGBoost) Machine Learning Algorithm by Using Well Log Data. Bio Web of Conferences. 2024;89:09003.
17. Xu G, Gan J, Li J, Liu H, Wu W. A Strain Fitting Strategy to Eliminate the Impact of Measuring Points Failure in Longitudinal Bending Moment Identification. Journal of Marine Science and Engineering. 2023;11(12):2282.
18. Deng Y-H, Luo X-Q, Yan P, Zhang N-Y, Liu Y, Duan SB. Outcome Prediction for Acute Kidney Injury Among Hospitalized Children via eXtreme Gradient Boosting Algorithm. Scientific Reports. 2022;12(1).
19. Zhang P, Jia Y, Shang Y. Research and Application of XGBoost in Imbalanced Data. International Journal of Distributed Sensor Networks. 2022;18(6):155013292211069.
20. Li D, Zhu D, Tao T, Qu J. Power Generation Prediction for Photovoltaic System of Hose-Drawn Traveler Based on Machine Learning Models. Processes. 2023;12(1):39.
21. Sheng S, Huang Y. Predicting 90-Day Readmission for Patients With Heart Failure: A Machine Learning Approach Using XGBoost. 2022.
22. Lv C-X, An S, Qiao B, Wu W. Time Series Analysis of Hemorrhagic Fever With Renal Syndrome in Mainland China by Using an XGBoost Forecasting Model. BMC Infectious Diseases. 2021;21(1).
23. Yao S, Wu Q, Kang Q, Chen YW, Yi L. An Interpretable XGBoost-based Approach for Arctic Navigation Risk Assessment. Risk Analysis. 2023;44(2):459-76.
24. Lutz M. Programming python: " O'Reilly Media, Inc."; 2001.
25. Ferguson B, Matyszak MK, Esiri MM, Perry VH. Axonal damage in acute multiple sclerosis lesions. Brain: a journal of neurology. 1997;120(3):393-9.
26. You Y, Barnett MH, Yiannikas C, Parratt J, Matthews J, Graham SL, et al. Chronic demyelination exacerbates neuroaxonal loss in patients with MS with unilateral optic neuritis. Neurology-Neuroimmunology Neuroinflammation. 2020;7(3).
27. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Annals of neurology. 2011;69(2):292-302.
28. Zeydan B, Atkinson EJ, Weis DM, Smith CY, Rocca LG, Rocca WA, et al. Reproductive History and Progressive Multiple Sclerosis Risk in Women. Brain Communications. 2020;2(2).
29. Brusola GA, Mitchell K, Ellison J, Gleeson P. Reliability and Validity of the Modified Walking and Remembering Test in Persons With Multiple Sclerosis. International Journal of Rehabilitation Research. 2021;44(4):323-9.
30. Mohamadirizi S, Shaygannejad V, Mohamadirizi S, Mohamadirizi M. Eating Disorders in a Multiple Sclerosis Clinical Population and Its Association With Social Anxiety. Journal of Multiple Sclerosis. 2016;03(03).
31. Brenton JN, Woolbright E, Briscoe-Abath C, Qureshi A, Conaway MR, Goldman MD. Body Mass Index Trajectories in Pediatric Multiple Sclerosis. Developmental Medicine & Child Neurology. 2019;61(11):1289-94.

32. Yalın OÖ, Edgünlü TG, Çelik S, Emre U, Güneş T, Erdal Y, et al. Novel SNARE Complex Polymorphisms Associated With Multiple Sclerosis: Signs of Synaptopathy in Multiple Sclerosis. Balkan Medical Journal. 2019;36(3):174-8.
33. Mele AA, Ogbuagu H, Parag S, Pierce RB. A Case of Multiple Sclerosis Uncovered Following Moderna SARS-CoV-2 Vaccination. Cureus. 2022.
34. Herrera-Acosta E, Garriga-Martina GG, Pérez JAS, Martínez-García E, Herrera-Ceballos E. Ixekizumab for Patients With Plaque Psoriasis Affected by Multiple Sclerosis. Sultan Qaboos University Medical Journal [Squmj]. 2021;21(3):488-90.
35. Ben-Fredj N. Lack of Correlation Between CCL5 -28c/G Functional Polymorphism and Multiple Sclerosis in Tunisian Patients. Journal of Clinical & Cellular Immunology. 2012;03(03).
36. Buchter B, Dunkel M, Li J. Multiple Sclerosis: A Disease of Affluence? Neuroepidemiology. 2012;39(1):51-6.
37. Kostoglou-Athanasiou I, Athanassiou L, Athanassiou P, Giannakopoulos A, Shoenfeld Y. Osteoporosis in a Woman With Multiple Sclerosis: A Case Report. Cureus. 2024.
38. Maloh DIA, AlNahar H, Maloh HIA. Quality of Life Among Multiple Sclerosis Patients in Terms of Mental Health. Asian Social Science. 2021;17(5):62.
39. Arsovski D. Assessment of Sexual Life and Social Life in People With Multiple Sclerosis. Health Promotion & Physical Activity. 2023;21(4):38-45.
40. Yalın OÖ, Edgünlü TG, Çelik S, Emre U, Güneş T, Erdal Y, et al. Novel SNARE Complex Polymorphisms Associated With Multiple Sclerosis: Signs of Synapthopathy in Multiple Sclerosis. Balkan Medical Journal. 2018.
41. Carbonell CF, Benson L, Rintell D, Prince JB, Chitnis T. Functional Relapses in Pediatric Multiple Sclerosis. Journal of Child Neurology. 2013;29(7):943-6.
42. Florea A, Maurey H, Sauter ML, Bellesme C, Sevin C, Deiva K. Fatigue, Depression, and Quality of Life in Children With Multiple Sclerosis: A Comparative Study With Other Demyelinating Diseases. Developmental Medicine & Child Neurology. 2019;62(2):241-4.