International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Interpretable Predictive Models for Alzheimer's Disease Using Explainable Boosting Machines

[1.]Ms.A. KAMATCHI, [2].Dr. V. MANIRAJ

[1] Research Scholar, Department of Computer Science,
A.V.V.M.Sri Pushpam College(Autonomous),Poondi,Thanjavur(Dt),Affiliated to Bharathidasan University,Thiruchirappalli,Tamilnadu
(Mail Id- kamatchia06@gmail.com)
[2] Associate Professor, Research Supervisor, Head of the Department,
Department of Computer Science,
A.V.V.M.Sri Pushpam College(Autonomous),Poondi,Thanjavur(Dt),Affiliated to Bharathidasan University,Thiruchirappalli,Tamilnadu
(Mail Id- manirajv61@gmail.com)

ABSTRACT :

Alzheimer's Disease (AD), a progressive neurological condition, presents an escalating health crisis globally. Timely detection is critical for effective intervention and slowing the disease's advancement. While advanced deep learning models have demonstrated superior performance in AD prediction, their complex and opaque nature often hinders clinical decision-making. This paper introduces a machine learning (ML) model that ensures accurate AD prediction while maintaining interpretability. Feature selection methods are utilized to enhance the relevance of the data features with respect to the target labels. Among various glass-box and black-box models examined for AD prognosis, the Explainable Boosting Machine (EBM) with Chi-square feature selection yields more accurate and interpretable outcomes, even with smaller datasets. The interpretability graphs provided by EBM offer both global and local explanations for the predictions, highlighting the key features that influence classification decisions. By providing transparency, EBMs enable greater trust in the model's decision-making, empowering healthcare professionals to make more informed and confident choices in AD diagnosis and treatment.

Keywords: Alzheimer's Disease, Black-box, Explainability, Explainable Boosting Machine, Glass-box, Feature Selection

1.INTRODUCTION :

Alzheimer's Disease (AD) is a progressive condition that causes the brain to shrink and damage brain cells, leading to the loss of memory and cognitive functions. AD is responsible for 60%-70% of dementia cases and is one of the primary causes of dementia. This disease is becoming an increasingly serious health issue, affecting millions of individuals, and the number of cases is expected to rise significantly in the future. AD is difficult to predict, and early diagnosis plays a crucial role in ensuring timely treatment to slow its progression. Traditional diagnostic methods, such as cognitive assessments, are often costly, time-consuming, and subjective, and they may fail to identify early-stage signs of the disease. The absence of a clear biomarker for AD further complicates diagnosis. Artificial Intelligence (AI) has emerged as a powerful tool in predicting various diseases, including AD. By analyzing large datasets, including medical records, brain scans, and genetic information, AI can identify patterns and estimate the risk of AD. Machine Learning (ML) methods such as Linear Regression (LR), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Decision Trees (DT) are straightforward models, though their accuracy is often limited. Neural Networks (NN) and other complex models tend to perform better but are difficult to interpret and computationally intensive. Ensemble models like XGBoost (XGB) and Gradient Boosting deliver improved performance but come with their own challenges in terms of interpretability.

Many Alzheimer's Disease (AD) detection models that demonstrate high accuracy rely on Deep Learning (DL) algorithms. However, DL models are highly dependent on large datasets to achieve precise results, and obtaining such extensive datasets in the medical field is a significant challenge. Additionally, balancing accuracy with interpretability remains a difficult issue. While simpler models like Linear Regression (LR) and Decision Trees (DT) are interpretable, they tend to offer lower accuracy. On the other hand, more complex models, such as ensemble methods and Neural Networks (NN), provide better accuracy but at the cost of reduced interpretability. This lack of transparency can undermine trust in these models and limit their practical application in clinical settings. For clinicians to confidently use these models for patient care, they must understand the factors that influence predictions. Moreover, datasets can vary in size, and with a large number of features, it becomes challenging for machine learning models to make accurate predictions. To address this, feature selection techniques can be utilized to automatically identify the most relevant features from the extensive set of input data.

The goal of this paper is to develop a model that not only provides high performance in terms of accuracy but also ensures transparency in its prediction process, using fewer features. In this context, the paper proposes the use of an Explainable Boosting Machine (EBM) to predict Alzheimer's Disease (AD) at an early stage, focusing on both accuracy and interpretability. EBM is a machine learning model that combines the benefits of boosting with the advantage of interpretability. Unlike complex deep learning models, the emphasis here is on creating simpler yet effective approaches. By incorporating feature selection techniques, the model improves prediction accuracy while maintaining clear decision-making transparency. The gradient boosting method refines the model by reducing the loss function. One of the key advantages of EBM is its ability to provide insights into how the model arrives at its predictions, with visual explanations of the results. This interpretability and flexibility make EBM especially useful for addressing complex machine learning problems where transparency and clarity are crucial. This approach offers several key benefits:

1. 1.**Increased Trust**: Clinicians require transparency to trust and adopt machine learning predictions in patient care. EBM fosters trust by providing explanations and identifying the factors that most influence the model's predictions.
2. 2.**Bias Mitigation**: Understanding the reasoning behind predictions helps to detect and mitigate biases within both the data and the model.
3. 3.**New Insights**: Recognizing the factors that influence the model's predictions opens opportunities for novel research and further exploration in the field.

This paper aims to contribute to the creation of a reliable, transparent, and accurate model that enhances patient care and leads to better outcomes by utilizing the power of the Explainable Boosting Machine (EBM). It also integrates feature selection techniques, such as chi-square and L1 regularization, to identify the most important features that influence the model's prediction of Alzheimer's Disease (AD) risk. The focus of this research is on using existing datasets and feature selection methods. While this study has been evaluated at a research level, it has not been validated through clinical trials and does not explore the development of new biomarkers for AD. To the best of our knowledge, this is the first work to apply EBM using the OASIS brain MRI dataset for Alzheimer's prediction. The key contributions of this paper are as follows:

1. 1.Explores various categories of machine learning models and proposes the most effective model for accurately predicting Alzheimer's Disease (AD).
2. 2.Utilizes feature selection techniques to identify the most relevant features, thus enhancing the model's performance accuracy.
3. 3.Examines multiple interpretable models to provide a better understanding of the predicted results.
4. 4.Extracts insights and generates explanations from the outcomes of the explainable model.
5. 5.Addresses the traditional trade-off between accuracy and interpretability.

The structure of the paper is as follows: Section II reviews the existing literature on the topic, Section III discusses the motivation for this work and provides background information. Section IV outlines the proposed predictive model and examines the different machine learning algorithms used. Section V presents the experiments, results, and discussion. Finally, the paper concludes in Section VI.

2.LITERATURE REVIEW :

Numerous studies have been conducted to predict Alzheimer's Disease (AD) using various machine learning (ML) and deep learning (DL) algorithms. However, relatively few researchers have explored the use of Explainable Boosting Machines (EBM) for this purpose. Some recent studies that applied the EBM model in different domains are reviewed here. In 2021, a study predicted AD based on MRI data from hippocampal subfields using EBM [3]. The study involved 200 brain MRIs from patients with Mild Cognitive Impairment (MCI), which were divided into two groups: progressive (pMCI) and stable (sMCI). The researchers achieved prediction accuracies of 80.5% and 84.2% for EBM without and with pairwise interactions, respectively.

In 2020, Harsha Nori and colleagues experimented with multiple classification and regression datasets using Explainable Boosting Machines (EBM) and incorporated Differential Privacy (DP) into their model [4]. The key advantages of their DP-EBM model included: (i) high accuracy, (ii) strong differential privacy, (iii) both global and local interpretability, and (iv) the ability to edit models without compromising privacy after training. One issue with machine learning models is that flawed datasets can lead the model to learn incorrect or biased patterns, resulting in inaccurate classifications or predictions. These flaws in the dataset are often not easy to identify.

Pathologic complete response (pCR) is a key factor in determining whether a patient with rectal cancer (RC) should undergo surgery following neoadjuvant chemoradiotherapy (nCRT). A model developed by [6] utilized Explainable Boosting Machines (EBM) to predict pCR in RC patients after nCRT, reducing the need for pathologist involvement in evaluating and determining pCR. In a similar vein, [7] applied the interpretable EBM classifier to predict pest presence in agricultural settings. This model incorporated factors like insect traps, weather forecasts, and vegetation indices to predict the onset of bollworm damage in cotton fields in Greece. Additionally, [8] proposed an explainable machine learning model to predict the compressive strength of concrete under varying mix ratios. Their research demonstrated how different raw material combinations impact concrete strength, enabling the identification of which mix ratios most influence compressive performance.

3.PRELIMINARIES :

3.1 Motivation

Deep neural networks and other advanced machine learning algorithms are widely used by developers to tackle complex, real-time problems. These models offer the significant advantage of achieving high performance and accuracy. However, they tend to be opaque, meaning end users cannot easily understand why a specific prediction was made, which factors influenced it, or how to alter the prediction. This lack of transparency, often referred to as the "black box" nature of these models, can create challenges in trust and decision-making. Relying solely on classification accuracy is not sufficient for critical decisions. This highlights the need for model explainability.

In many domains, it is essential for a model to explain the reasoning behind its predictions. For instance, in healthcare, a patient has the right to understand why their MRI result was classified as showing a tumor. Similarly, an applicant should know the reasons behind their loan rejection, and a court cannot simply base a prison sentence on the results of an AI model. These situations necessitate transparency in how the model arrived at its decision. Providing such explanations enhances user trust and ensures accountability.

Explanations are vital for ensuring fairness, privacy, reliability, and trust in machine learning models. However, balancing accuracy and interpretability is often a challenge, as achieving both high accuracy and transparency in a model typically requires compromises.

3.2 Explainability

Explainability, or interpretability, refers to the degree to which a human can understand a model's predictions and behavior. Machine learning models, particularly complex ones, often function as black boxes and do not offer easily understandable explanations. A model must be able to provide explanations both at the global level, where it outlines its general behavior and the significant factors it considers, and at the local level, where it explains individual predictions. Global explanations provide insight into the model's priorities and can uncover any flaws in decision-making, while local explanations offer a detailed account of how a specific prediction was made.

Simple, inherently interpretable models like Logistic Regression (LR) and Decision Trees (DT) are more transparent, as their predictions are easier for users to understand. However, these models often do not perform as well as more complex algorithms. To address this, developers apply interpretability techniques to existing complex models. This process, known as model-agnostic interpretation, separates the explanation from the underlying model itself, allowing users to understand complex predictions without compromising the model's effectiveness.

3.3 Glassbox v/s Whitebox Models

ML algorithms falls under either glasssbox or blackbox model.

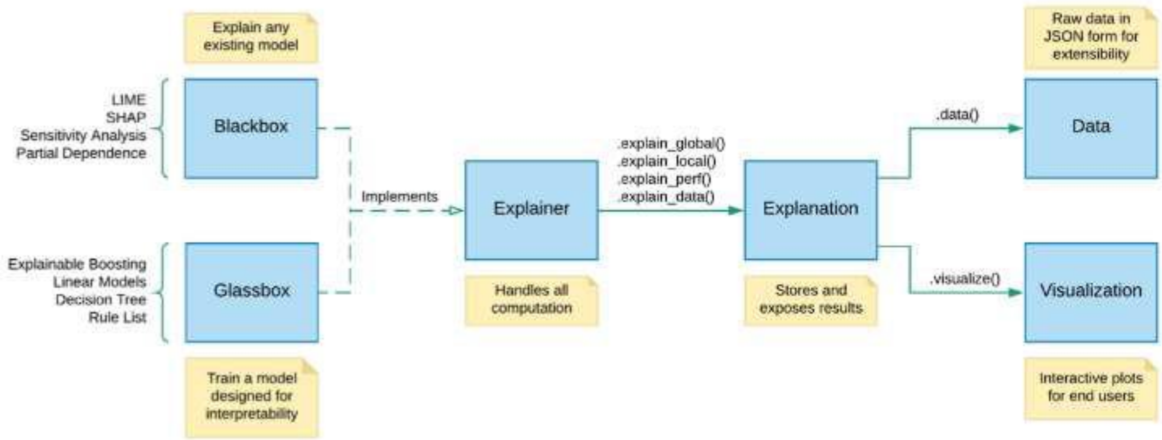
3.3.1 Glassbox models

Glassbox models are designed to be fully interpretable, often achieving accuracy levels comparable to more advanced methods. These models are capable of providing clear explanations at both the global and local levels. In the case of glassbox models, these explanations are exact, meaning they provide a clear and accurate understanding of the decision-making process the model used to reach its prediction. Such explanations are useful in helping end users understand which factors had the most significant influence on a particular prediction. Logistic Regression (LR) and Decision Trees (DT) are examples of glassbox models.

3.3.2 Blackbox models

Black-box models are known for their ability to produce predictions or make decisions without revealing the processes or methodologies behind them. The internal workings, including the specific factors and their associated weights used in these models, are hidden from the user. This lack of visibility creates a significant barrier to understanding, resulting in limited transparency. Interpretability techniques for black-box models aim to extract explanations from any machine learning algorithm or framework. These techniques are particularly useful in complex workflows where not all components are directly interpretable, such as in model ensembles, pre-processing stages, and sophisticated models like deep neural networks. Figure 1 illustrates various approaches to explaining AI models.

InterpretML is an open-source tool that includes state-of-the-art techniques for model interpretability . It provides two types of interpretability: glassbox and black-box, and also features a visualization platform .



3.3.3 ML models employed in the study

The ML models implemented and analyzed in this experiment are explained below.

3.3.3.1 Logistic Regression

It is a simple and well analysed model. When the dependent variables are categorical, logistic regression is used.

3.3.3.2 Decision Tree

Decision trees are another type of straightforward and interpretable model. As non-parametric supervised algorithms, they visually represent the decision-making process. Decision trees are easy to understand and interpret, making them an accessible choice for many applications [18].

3.3.3.3 XGBoost

This algorithm is an implementation of optimized gradient-boosted decision trees, designed with a focus on speed and performance. It falls under ensembling techniques, where the predictions from multiple weak learners are combined to generate a stronger, more accurate prediction. Additionally, it can effectively handle missing data .

3.3.3.4 ExtraTree Classifier

It refers to the Extremely Randomized Tree Classifier, an ensembling algorithm that combines the results of multiple decision trees. Each tree is constructed using the original dataset, and for each node, a random subset of k features from the feature set is selected. A mathematical criterion, such as the Gini Index or Information Gain, is then used to identify the best feature for splitting the node, effectively dividing the samples into two distinct groups.

3.3.3.5 Voting Classifier

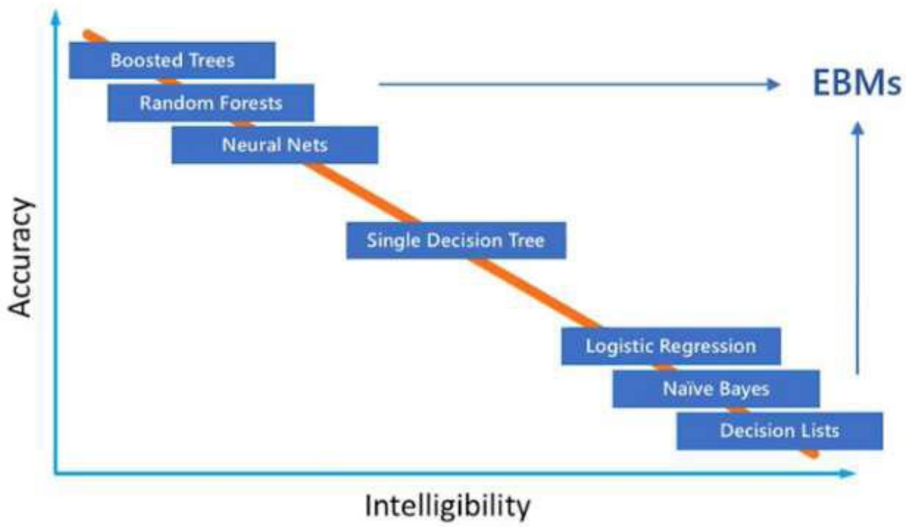
In this context, various different ML algorithms are integrated and employed to predict class labels by either taking a majority vote (hard voting) or considering the average predicted probabilities (soft voting).

3.3.3.6 Multi-Layer Perceptron

An MLP (Multi-Layer Perceptron) is a type of fully connected, feedforward neural network distinguished by its structure, which consists of an input layer, an output layer, and one or more hidden layers in between. MLP is designed to address supervised learning tasks.

3.3.3.7 Explainable Boosting Machine

The EBM (Explainable Boosting Machine), developed by Microsoft, is an interpretable model based on the boosting technique. There is often a trade-off between the accuracy and interpretability of a model. Simple glass-box models are typically more interpretable but tend to have lower accuracy, whereas complex black-box models like neural networks provide better predictive performance but lack transparency. Black-box explainers are used to provide surface-level explanations for these opaque models. In contrast, glass-box models are inherently interpretable by design. The key advantage of EBM is that it balances strong performance with the ability to deliver clear, informative explanations. Figure 2 illustrates the performance-interpretability trade-off of EBM compared to other machine learning algorithms.



The graph above can be interpreted in the following way: Simple models like Logistic Regression (LR) and Decision Trees (DT) generally lack high accuracy. On the other hand, complex black-box models such as Neural Networks provide good performance but are difficult to interpret. Ensemble models like Random Forests and boosted trees offer better performance, yet they are also challenging to interpret. EBMs occupy the top-right corner of the graph, which is a positive outcome for developers. This placement highlights that EBMs enable the creation of models that are both highly accurate and easy to interpret.

3.4 Feature Selection techniques

Feature selection techniques in machine learning are used to identify the most relevant features for building an optimized model. By applying these techniques, model accuracy can be enhanced, while also reducing training time and minimizing overfitting [26]. The most widely used methods for supervised feature selection include filters, wrappers, embedded methods, and hybrid approaches [2]. Each of these methods offers a variety of techniques for selecting features. The feature selection methods implemented in this study are:

3.4.1 Chi-Square Method

This technique falls under the filter method, where features are chosen based on their chi-square score, independent of the machine learning model used. It applies a straightforward statistical test to the categorical features in the dataset. The chi-square test assesses the independence of each feature from the target variable and determines the significance of each feature in relation to the target.

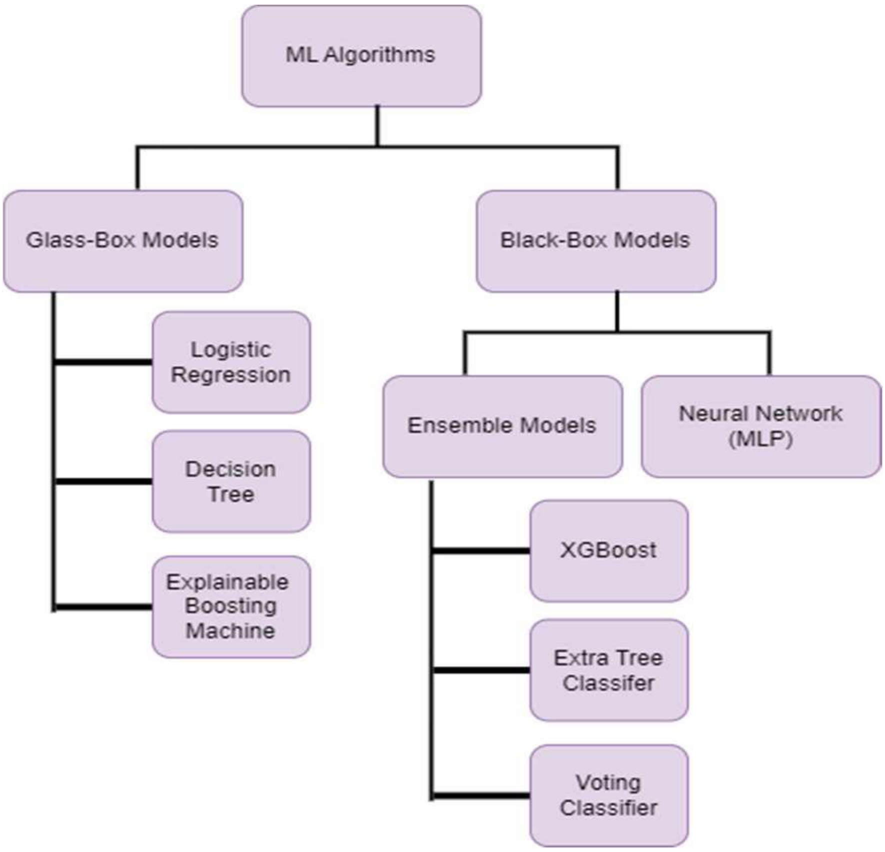
LASSO Regularization (L1) Regularization introduces a penalty term to the parameters of a machine learning model to help prevent overfitting and enhance generalization. L1 regularization, in particular, can shrink some coefficients to zero, effectively eliminating redundant or irrelevant features and preventing the model from becoming overly complex. This technique is an embedded method, offering advantages of both filter and wrapper approaches.

4. METHODOLOGY :

4.1 Overview

This research employs a ML model with explainability design o detect the Ad in an accurate and interpretable way. The data collected form OASIS dataset contains demographic information of human beings, cognitive test scores, brain scan measurements etc. As a next step, clean and preprocess the data by handling missing values, outliers and scaling numerical features. And then employ feature selection techniques such as chi-square and L1 regularization for identifying relevant information contributing to the AD risk prediction. Choose EBM and train the model and obtain the feature importance scores. Evaluate the model performance using standard metrics, analyze the feature importances and interpret the results globally and locally. This study proposes different categories of ML algorithms for the prediction of AD. Figure 4 shows the ML algorithms employed for model training. They are as follows:

- Simple and by design interpretable models, which comes under the category of Glassbox models, such as LR, DT, and EBM.
- As ensemble techniques and neural networks produces greater accuracy in general, these models are also employed in this experiment. But these Black-box models are not interpretable. XGBoost, Extra Tree Classifier and Voting Classifier are used as Ensemble techniques.
- MLP is used to build a basic Neural Network.



4.2 Dataset Description

The Alzheimer's dataset includes longitudinal brain MRI data from the Open Access Series of Imaging Studies (OASIS). OASIS is a publicly available collection of neuroimaging datasets intended for research use. The dataset comprises data from 150 subjects and includes 14 key attributes, such as age, socioeconomic status (SES), years of education, gender, and various MRI-related features like MMSE, CDR, eTIV, nWBV, and ASF . These features are essential for the classification task. Each subject is categorized as either demented or non-demented.

4.3 The Proposed Predictive Model

The proposed framework is outlined in Figure 5. The longitudinal MRI dataset from OASIS [27] is used to predict Alzheimer's Disease (AD). The data undergoes preprocessing, including handling missing values and encoding categorical variables. Feature selection techniques, such as Chi-square and L1 regularization, are applied to the categorical data. Next, the dataset is divided into training and testing sets. Various machine learning algorithms are employed for training, including Glass-box models like Logistic Regression (LR), Decision Trees (DT), and Explainable Boosting Machines (EBM), as well as Black-box models such as XGBoost, Random Forest (RF), ExtraTree Classifier, Voting Classifier, and Multi-Layer Perceptron (MLP). Both Glass-box and Black-box models are implemented, with EBM outperforming the others in terms of both accuracy and interpretability. EBM is a tree-based model that integrates cyclic gradient boosting with Generalized Additive Model principles, automatically detecting interactions between variables [24]. Notably, EBM achieves accuracy levels similar to state-of-the-art black-box models, such as XGBoost and LightGBM, while offering complete interpretability. Therefore, EBM is selected for the framework, providing both improved accuracy and transparency in AD prediction.

5. EXPERIMENTS, RESULTS & DISCUSSIONS :

To assess and compare the effectiveness of EBM, some of the glassbox and blackbox algorithms are implemented here in OASIS MRI Brain image dataset for Alzheimer's prediction.

5.1 Experiments

Algorithms from different categories are employed to compute the efficiency of the proposed model.

5.1.1 Logistic Regression

Since this model is simple, the results are also interpretable. It had shown the least accuracy of 60% without any feature selection and 65% in Chi-square technique.

5.1.2 Decision Tree

This is another by design explainable model. This algorithm generated a performance accuracy of 71% and 75% without feature selection and with Chisquare feature selection respectively.

5.1.3 XGBoost

A boosting technique which is extensively used by the researchers in different domains produced an accuracy of 86% and 89% respectively without feature selection and with Chi-square feature selection respectively.

5.1.4 ExtraTree Classifier

This ensemble classifier produced an accuracy of 89% and 90% respectively without and with feature selection techniques.

5.1.5 Voting Classifier

This model generated an accuracy of 90% without feature selection and 92% with Chi-square feature selection, which is the second highest performance accuracy among all other models.

5.1.6 Multi Layer Perceptron

A Simple Neural Network is also employed in this experiment. It could produce only a low accuracy of 61% without feature selection but 84% with Chi-square feature selection which is a greater improvement.

5.1.7 Explainable Boosting Machine

This model is by design interpretable and produced the highest performance accuracy 93% without any feature selection techniques and 95% and 94% respectively with Chi-square and L1 feature selection techniques respectively. The major highlight of this model is that the results of this model are explainable. It could generate both global and local explanations.

5.2 Comparison of the proposed model with existing models

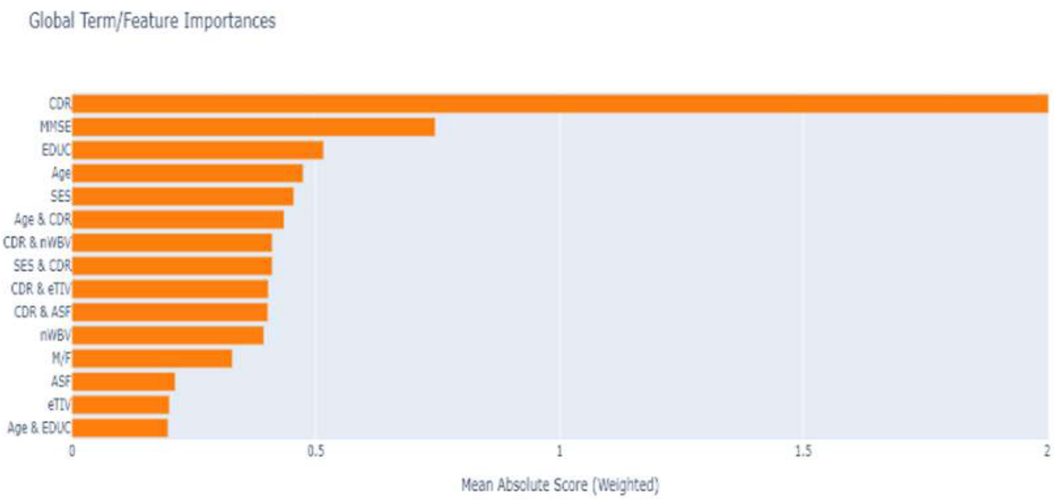
. It is evident from the table that the proposed EBM model outperforms the other models in terms of performance accuracy. The data clearly demonstrates that the EBM model is more effective at classifying AD compared to the other existing models reviewed in the passage.

5.3 Interpretability Graphs

OASIS Brain MRI dataset is considered for the illustration of Interpretability graphs.

5.3.1 Global Explanations

This explains the overall contribution of features to the model. Contribution of individual features and its relation to the model would be clearly shown here. These explanations shows us what a model finds important. It also helps us to identify the potential flaws of a model in decision making. The global importance of each feature is calculated by taking the mean absolute contribution of a feature or its interaction to the predictions averaged across the training dataset.



6.CONCLUSION :

Accurately and interpretable predicting a disease is a complex challenge. This study proposes a novel approach for Alzheimer's Disease (AD) detection using Explainable Boosting Machines (EBM). The model shows promising results in accurately predicting AD risk while providing clear, interpretable insights into the factors that contribute to these predictions. This emphasis on explainability not only builds trust but also facilitates the integration of the model into clinical practices. Both Glass-box models, including Logistic Regression, Decision Trees (DT), and EBM, as well as black-box models like XGBoost, Extra Tree Classifier, Voting Classifier, and Multi-Layer Perceptron, were used in the experiment. By incorporating an effective feature selection method, the Chi-square technique, the model's performance accuracy was further enhanced. The results indicate that EBM, combined with Chi-square feature selection, outperformed other models in prediction accuracy while offering greater transparency and providing valuable insights into the predictions. This model has the potential to improve AD diagnosis, as the results are self-explanatory to clinicians, helping them make confident decisions. The intelligibility provided by the EBM graphs fosters the creation of new knowledge and aids in better understanding and debugging the data. This paper addresses the performance-interpretability trade-off in machine learning models. The contribution of this work lies in advancing the field of explainable AI (XAI) for AD detection. By offering a transparent and interpretable model, this research can transform early AD diagnosis, leading to timely interventions and better patient outcomes.

7.REFERENCES :

[1] World Health Organization. (n.d.). Mental health: Neurological disorders. World Health Organization. https://www.who.int/newsroom/ questions-and-answers/item/mentalhealth- neurological-disorders (Accessed Jan 16, 2024)

[2] A. Jović, K. Brkić and N. Bogunović, "A review of feature selection methods with applications," 2015 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2015, pp. 1200- 1205, doi: 10.1109/MIPRO.2015.7160458.

[3] Quattrone, A., Quattrone, A. (2021). Explainable Boosting Machine for Predicting Alzheimer's Disease from MRI Hippocampal Subfields. In: Mahmud, M., Kaiser, M.S., Vassanelli, S., Dai, Q., Zhong, N. (eds) Brain Informatics. BI 2021. Lecture Notes in Computer Science(), vol 12960. Springer, Cham. https://doi.org/10.1007/978-3-030- 86993-9_31.

[4] Nori, H., Caruana, R., Bu, Z., Shen, J.H. & Kulkarni, J.. (2021). Accuracy, Interpretability, and Differential Privacy via Explainable Boosting. Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research, 139:8227-8237 Available from ttps://proceedings.mlr.press/v139/nori21a.html.

[5] Chen, Zhi & Tan, Sarah & Nori, Harsha & Inkpen, Kori & Lou, Yin & Caruana, Rich. (2021). Using Explainable Boosting Machines (EBMs) to Detect Common Flaws in Data. 10.1007/978-3-030-93736-2_40.

[6] Wang Du, Lee Sang Ho, Geng Huaizhi, Zhong Haoyu & et al, Interpretable machine learning for predicting pathologic complete response in patients treated with chemoradiation therapy for rectal adenocarcinoma, Frontiers in Artificial Intelligence, Vol. 5, 2022, https://www.frontiersin.org/articles/10.3389/fra i.2022.1059033, doi. 10.3389/frai.2022.1059033, ISSN=2624-8212

[7] O. Nanushi, V. Sitokonstantinou, I. Tsoumas and C. Kontoes, "Pest Presence Prediction Using Interpretable Machine Learning," 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), Nafplio, Greece, 2022, pp. 1-5, doi: 10.1109/IVMSP54334.2022.9816284.

[8] Gaoyang Liu, Bochao Sun, Concrete compressive strength prediction using an explainable boosting machine model, Case Studies in Construction Materials, Volume 18, 2023, e01845, ISSN 2214-5095, https://doi.org/10.1016/j.cscm.2023.e01845.

[9] Değer, Zeynep & Taskin, Gulsen & Wallace, John. (2023). Estimate Deformation Capacity of Non-Ductile RC Shear Walls using Explainable Boosting Machine. 10.48550/arXiv.2301.04652.

[10] Sarica, A., Quattrone, A., Quattrone, A. (2022). Introducing the Rank-Biased Overlap as Similarity Measure for Feature Importance in Explainable Machine Learning: A Case Study on Parkinson's Disease. In: Mahmud, M., He, J., Vassanelli, S., van Zundert, A., Zhong, N. (eds) Brain Informatics. BI 2022. Lecture Notes in Computer Science(), vol 13406. Springer, Cham. https://doi.org/10.1007/978-3-031- 15037-1_11

[11] Maxwell, A. E., Sharma, M., & Donaldson, K. A. (2021). Explainable boosting machines for slope failure spatial predictive modeling. Remote Sensing, 13(24), 4991. https://doi.org/10.3390/rs13244991

[12] Bosschieter, Tomas & Xu, Zifei & Lan, Hui & Lengerich, Benjamin & Nori, Harsha & Sitcov, Kristin & Souter, Vivienne & Caruana, Rich. (2022). Using Interpretable Machine Learning to Predict Maternal and Fetal Outcomes. 10.48550/arXiv.2207.05322.

[13] Archana P. Menon, Dr. R. Gunasundari, Study of Interpretability in ML Algorithms for Disease Prognosis, Revista Geintec-Gestao, Inovacao e Technologias, Vol. 11 No. 4 Aug 2021, Pages 4735-4749, ISSN NO: 2237-0722. Available: https://www.revistageintec.net/index.php/revist a/article/view/2500

[14] O. Conor Sullivan, Interpretability in Machine Learning, Towards Data science, Oct.2020. Accessed on: Jun. 12, 2021. [Online]. Available: https://towardsdatascience.com/interpretabilityin- machine-learning-ab0cf2e66e1

[15] Molnar, Christoph. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable, 2019. https://christophm.github.io/interpretable-mlbook/.

[16] Nori, Harsha and Jenkins, Samuel and Koch, Paul and Caruana, Rich, "InterpretML: A Unified Framework for Machine Learning Interpretability." ArXiv abs/1909.09223 (2019)

[17] Julien I.E. Hoffman, Chapter 33 – Logistic Regression, Basic Biostatistics for Medical and Biomedical Practitioners (Second Edition), Academic Press, 2019, Pages 581-589, ISBN 9780128170847, https://doi.org/10.1016/B978- 0-12-817084-7.00033-4

[18] 1.10. Decision Trees. (n.d.). Scikit-learn. https://scikit-learn.org/stable/modules/tree.html (Accessed Jan 16, 2024)

[19] GeeksforGeeks. (2023, February 6). XGBoost. GeeksforGeeks, https://www.geeksforgeeks.org/xgboost/ (Accessed Jan 16, 2024)

[20] Sklearn.tree.extratreeclassifier. scikit. (n.d.). https://scikitlearn. org/stable/modules/generated/sklearn.tree .ExtraTreeClassifier.html (Accessed Jan 16, 2024)

[21] Sklearn.ensemble.VotingClassifier. scikit. (n.d.). https://scikitlearn. org/stable/modules/generated/sklearn.ens emble.VotingClassifier.html (Accessed Jan 16, 2024)

[22] 1.17. neural network models (supervised). scikit. (n.d.). https://scikitlearn. org/stable/modules/neural_networks_sup ervised.html (Accessed Jan 16, 2024)

[23] YouTube. (2020, May 16). The science behind InterpretML: Explainable boosting machine. YouTube. https://www.youtube.com/watch?v=MREiHgH gl0k (Accessed April 16, 2023)

[24] Explainable boosting machine — interpretml documentation. https://interpret.ml/docs/ebm.html, (Accessed Jan 16, 2024)

[25] Lou, Y., Caruana, R., Gehrke, J. & Hooker, G. (2013). Accurate intelligible models with pairwise interactions. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining – KDD '13: ACM Press.

[26] S. Picek, A. Heuser, A. Jovic and L. Batina, "A Systematic Evaluation of Profiling Through Focused Feature Selection," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 27, no. 12, pp. 2802-2815, Dec. 2019, doi: 10.1109/TVLSI.2019.2937365.