

A Novel Approach to Dementia Prediction of DTI Markers Using BALI, LIBRA, and Machine Learning Techniques

Ahmad Akbarifar¹, Adel Maghsoudpour¹, Fatemeh Mohammadian², Morteza Mohammadzaheri³, Omid Ghaemi⁴

¹Department of Mechanical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

²Department of Psychiatry, Roozbeh Hospital, Tehran University of Medical Sciences, Tehran, Iran

³College of Engineering, Birmingham City University, Birmingham, UK

⁴Department of Radiology and Interventional Radiology, Imam Khomeini Hospital (Imaging Centre) and Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran

A. A. is a PhD candidate at the Department of Mechanical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran (e-mail: ahmad.akbarifar@srbiau.ac.ir).

A. M. is an Assistant Professor of the Department of Mechanical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran (corresponding author to provide phone: 021-44865154-8; fax: 021-44865166; Postal Code 1477893855Tehran, Iran; e-mail: a.maghsoudpour@srbiau.ac.ir).

F. M. is an Assistant Professor at the Department of Psychiatry, Roozbeh Hospital, Tehran University of Medical Sciences, Tehran, Iran (e-mail: fmohammadianr@sina.tums.ac.ir).

M. M. is a Senior Lecturer at the College of Engineering, Birmingham City University, Birmingham, B4 7XG, UK (e-mail: morteza.mohammadzaheri@bcu.ac.uk).

O. G. is an Assistant Professor at the Department of Radiology and Interventional Radiology, Imam Khomeini Hospital (Imaging Centre) and Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran (e-mail: oghaemi@sina.tums.ac.ir).

Conflicts of Interest

All authors declare that they have no conflicts of interest to disclose.

Consent Statement

All authors confirm that all human subjects, involved in the reported research, provided informed consent.

Competing interests

The authors declare no competing interests.

Acknowledgment

The authors wish to especially thank at Advanced Diagnostic and Interventional Radiology Research Centre (ADIR), Imam Khomeini Complex Hospital, Tehran University of Medical Science, Tehran, Iran, for their kind support of this project.

Abstract

Early prediction of dementia and disease progression remains challenging. This study presents a novel machine learning framework for dementia diagnosis by integrating multimodal neuroimaging biomarkers and inexpensive and readily available clinical factors. Fractional anisotropy (FA) measurements in diffusion tensor imaging (DTI) provide microstructural insight into white matter integrity disturbances in dementia. However, the acquisition of DTI is costly and time-consuming. We applied Recursive Feature Elimination (RFE) to identify predictors from structural measures of the 9 factors of Brain Atrophy and Lesion Index (BALI) factors and 42 factors of Clinical Lifestyle for Brain Health (LIBRA) factors to estimate FA in DTI. The 10 most effective features of BALI/ LIBRA selected by RFE were used to train an interpretable decision tree model to predict the severity of dementia from DTI. A decision tree model based on biomarkers selected by RFE achieved an accuracy of 96.25% in predicting dementia in an independent test set. This integrated framework pioneers the prediction of white matter microstructural changes from available structural/ clinical factors using machine learning. By avoiding DTI acquisition, our approach provides a practical and objective tool to improve dementia screening and progress monitoring. The Identification of key predictive markers of BALI/ LIBRA will also provide information on the mechanisms of lifestyle-related disease mechanisms, neurodegeneration, and white matter dysfunction. This study aims to predict FA measures from DTI, which indicate white matter integrity and dementia severity, using inexpensive and readily available BALI and LIBRA factors through machine learning.

Keywords— *Recursive feature elimination, decision tree, prediction, dementia, MRI*

Introduction

Dementia represents an urgent public health crisis that is growing rapidly in scale and impact on society [1, 2]. Currently affecting more than 55 million people worldwide, the prevalence is expected to triple to 152 million by 2050 [3]. The associated economic burden is more than \$1 trillion per year. Beyond costs, dementia wreaks enormous psychosocial havoc on patients and families [4, 5]. As the population ages, this growing epidemic affects healthcare systems [6]. There is a strong need for solutions that can detect the onset of the disease to delay the progression of the disease with the necessary treatments [7].

Despite extensive research, the diagnosis of dementia remains challenging, especially in the early stages, when symptoms are subtle [8]. Current methods are heavily based on neuropsychological assessments, clinical observations, and structural imaging [9]. Although MRI and tomography can show severe atrophy, they lack the quantitative biomarkers needed to monitor subtle changes [10]. These limitations emphasize the need for more objective, sensitive and scalable diagnostic methods that can detect the underlying pathophysiological processes of dementia before overt symptoms appear [2, 11]. Recently, computational approaches in biomedical engineering are being effectively used to treat various patients [12, 13].

Advanced MRI techniques such as BALI and DTI provide quantitative biomarkers of neurodegeneration with the potential for earlier and more accurate diagnosis of dementia [14-16]. BALI enables the quantification of regional atrophy through segmentation and volumetrics [17]. DTI characterizes microstructural changes in white matter integrity through factors such as fractional anisotropy [18-20]. Unlike conventional imaging, these methods can detect pathophysiological changes that precede severe atrophy [10, 21]. Each provides more information about the complex etiology of dementia [22]. A research priority is to develop ways to integrate multimodal data to use this combined insight for diagnosis and monitoring [23].

The LIBRA method includes entries related to lifestyle factors that affect brain health. These inputs include diet, physical activity, cognitive engagement, social interaction, sleep quality, and other modifiable risk factors [24]. LIBRA's comprehensive assessment of a person's diet includes components such as fruit/ vegetable consumption, nutritional balance, and regularity of meals [25]. It also examines physical activity patterns, including aerobic exercise, strength training, and overall activity level [26]. Also, the inputs to cognitive engagement inputs reflect intellectual stimulation, new learning, and mental challenge [27]. The input of social interaction covers social contact, relationship quality, and community participation [28]. Inputs assess sleep quality, duration, stability, disturbances, and daytime sleepiness [29]. Other factors include cardiovascular health, drug and smoking use, stress level, and demographic factors. In total, LIBRA's 42 entries provide a multidimensional profile of lifestyle factors affecting brain health [30].

Many previous studies have evaluated the diagnostic utility of BALI, LIBRA, and DTI separately [30]. For example, several studies using BALI-derived volumetrics such as hippocampal atrophy have demonstrated accuracy in distinguishing Alzheimer's patients from controls [31]. Similarly, LIBRA-based quantification for the diagnosis of dementia and comorbidities has shown promise in past research [32]. DTI measurements have shown changes before atrophy in frontotemporal dementia, and extensive research has been done on its relationship with the diagnosis of dementia [33]. However, most studies focused on a single method, lacking the integration of complementary information that could improve diagnosis and predictive ability, especially in complex and heterogeneous conditions such as dementia.

Dementia has a complex multifactorial pathogenesis that involves various pathophysiological processes [34]. Unimodal approaches cannot understand this heterogeneity well and may miss some changes [35]. A multimodal approach that integrates complementary structural, microstructural, and vascular biomarkers from BALI, LIBRA, and DTI could enable a more comprehensive and accurate description of the pathophysiology of dementia for diagnosis and monitoring [36]. Machine learning is an ideal framework for integrating such heterogeneous data into sensitive predictive models [37]. Therefore, this study seeks to develop a machine learning model that combines multimodal neuroimaging biomarkers to enhance the early detection of dementia and predict disease progression.

Machine learning is a powerful approach to integrate diverse multivariate data and build predictive models [38]. Unlike traditional statistical methods, machine learning can model complex non-linear interactions between many variables [39]. Decision tree algorithms can integrate heterogeneous inputs such as neuroimaging biomarkers and clinical data to maximize diagnostic or predictive accuracy [40]. In addition, machine learning includes cross-validation and independent testing to protect against overfitting in limited samples [41, 42]. This allows robust generalization to new data [43]. Interpretability of algorithms, such as decision trees, also facilitates inference to key predictive biomarkers and relationships underlying dementia [44]. Therefore, machine learning was chosen as an optimal framework for combining multivariate data in a predictive model of dementia progression [45].

While numerous studies have independently explored the diagnostic utility of BALI, LIBRA, and DTI biomarkers for dementia, there remains a critical need to develop integrative frameworks that leverage the complementary strengths of these multimodal data sources [46]. Recent years have witnessed a growing recognition that the complex pathophysiology of dementia requires a holistic approach that combines structural, lifestyle, and microstructural neuroimaging markers [47]. Consequently, an emerging area of research has focused on machine learning techniques to synthesize these heterogeneous data modalities into predictive models of disease progression and severity [45, 48, 49]. Notable efforts include the use of support vector machines, random forests, and deep learning architectures to fuse volumetric, diffusion metrics, and clinical/ lifestyle factors derived from MRI for dementia classification and staging [50]. However, the integration of comprehensive BALI, LIBRA, and DTI assessments within a unified predictive framework remains an open challenge. In addition, there is a pressing need for approaches that can leverage readily available clinical data to estimate costly and time-intensive neuroimaging protocols such as DTI, thereby enabling more scalable and cost-effective screening paradigms. By developing a machine learning approach to predict white matter microstructural integrity from structural MRI and lifestyle

factors, the current study represents a novel contribution in this direction and does away with the necessity for widespread DTI acquisitions.

Furthermore, while previous studies have explored relationships between individual BALI/ LIBRA factors and neuroimaging metrics, there remains an outstanding need to elucidate the most salient predictors of white matter integrity from these multidimensional lifestyle and structural assessments [51]. By strategically reducing the dimensionality of the BALI/ LIBRA feature space through recursive feature elimination, the current work delineates a refined subset of factors that are maximally informative of the DTI measures involved in dementia pathology. This data-driven feature selection approach not only enhances predictive performance but also facilitates the interpretability of the machine learning model, providing insight into the key lifestyle and neuroanatomical determinants of white matter degeneration. Ultimately, the identification of these predictive relationships has the potential to advance our fundamental understanding of the multifactorial etiology of dementia, highlighting modifiable risk factors that could be targeted through early interventions or lifestyle modifications.

This study aimed to develop a machine learning model by integrating quantitative biomarkers from BALI, LIBRA, and DTI to predict dementia severity [52]. We hypothesized that combining multiple methods would allow more accurate staging of dementia progression compared to individual biomarkers alone [46]. This could ultimately revolutionize management by providing an objective, scalable, and cost-effective approach to predicting the progression of dementia from an early and treatable stage. Furthermore, the development of a model that can show the most effective factors that affect fractional anisotropy can help in this field and lead the way in preventing dementia [53].

The main contribution of this work is to leverage inexpensive and routinely collected brain imaging (BALI) and lifestyle/ clinical (LIBRA) factors to estimate costly and time-consuming measures of white matter integrity. This approach can potentially enable cost-effective screening and monitoring of dementia progression without requiring acquisition of DTI for all individuals.

Although BALI, LIBRA, and DTI have shown utility individually, a key novelty of our work is the integration of these complementary modalities into a unified machine learning framework. By combining structural brain imaging (BALI), lifestyle/ clinical factors (LIBRA), and microstructural white matter integrity DTI, the aim is to use their synergistic strengths for improved dementia prediction. Moreover, our approach explores using the relatively inexpensive and easily obtained BALI/ LIBRA factors to estimate the costly DTI measures, enabling more scalable screening and monitoring applications.

The primary motivation for this study stems from the urgent need to develop practical and cost-effective methods for early detection and monitoring of dementia progression. Although advanced neuroimaging techniques like DTI provide valuable information on microstructural integrity of white matter, their widespread implementation is hindered by the associated costs and complexity of data acquisition [54]. Using readily available structural brain imaging data from BALI and comprehensive lifestyle assessments from LIBRA, this study pioneers a novel approach to estimate markers derived from DTI of white matter health. Integration of multimodal data sources through machine learning techniques not only circumvents the need for costly DTI scans, but also enables a holistic characterization of the complex interplay between lifestyle factors, macroscale brain changes, and microstructural white matter degeneration in dementia. The recursive feature elimination strategy employed here uniquely identifies the most salient predictors of BALI and LIBRA of the DTI measures, elucidating key lifestyle and neuroanatomical determinants of white matter integrity [55]. This data-driven approach is of profound significance in advancing our fundamental understanding of the etiology of dementia, highlighting potential targets for early intervention and lifestyle modification. Ultimately, the proposed framework paves the way for a practical, scalable, and cost-effective paradigm for dementia screening and progression monitoring, addressing a critical unmet need in global healthcare.

Although prior studies have demonstrated the individual utility of the BALI, LIBRA, and DTI modalities for the assessment of dementia, our work represents a pioneering effort to synergistically integrate these complementary data sources within a unified predictive framework. In contrast to previous pairwise combinations or unimodal approaches, our multimodal machine learning model captures the multifaceted pathophysiology of dementia by fusing structural neuroimaging, comprehensive lifestyle profiling, and microstructural white matter characterization. A key novelty lies in our recursive feature elimination strategy to pinpoint the most predictive BALI and LIBRA factors of white matter integrity disruption, elucidating potential targets for preventive lifestyle interventions. Furthermore, the interpretable decision tree formulation offers potential for individualized risk stratification and personalized treatment planning by delineating multimodal decision rules related to disease burden [56]. Beyond diagnostic applications, the elucidated associations between modifiable lifestyle factors, macroscopic brain changes, and microstructural white matter alterations could stimulate new mechanistic investigations into the complex etiology of dementia and pathways for intervention.

Problem Statement

The present study expands on the use of clinical data to identify important BALI and LIBRA factors in estimating DTI measures of white matter integrity. Despite the recognized utility of DTI for assessing white matter health, the acquisition of DTI data is costly and time-intensive. In contrast, the numerous lifestyle indices of BALI and LIBRA and structural brain are the more easily achieved. Therefore, the elucidation of specific BALI/ LIBRA factors most predictive of key DTI metrics could enable a more practical evaluation of white matter status. Leveraging an expanded dataset with both BALI/ LIBRA and DTI metrics, the current work implements recursive feature elimination and decision tree methodologies to determine a subset of BALI/ LIBRA predictors that is maximally informative for DTI measures known to indicate dementia. By strategically reducing the dimensionality of the BALI/ LIBRA feature space, this study aims to refine the understanding of associations between lifestyle risk, macroscale brain health, and white matter microstructure, without requiring DTI acquisition for all individuals. The selected lifestyle and clinical indices may also facilitate the interpretation of previous models linking BALI/ LIBRA factors to DTI. In summary, the present work seeks to enhance the efficiency of evaluating the integrity of white matter integrity using readily obtained clinical data.

The present study expands on the use of clinical data from BALI and LIBRA factors to estimate FA measures from DTI, which are indicators of white matter integrity and severity of dementia. Despite the recognized utility of DTI for assessing white matter health, the acquisition of DTI data is costly and time-intensive. In contrast, the numerous BALI structural brain imaging factors and LIBRA lifestyle factors are more easily reached.

Materials and Methods

Data Collection

Data were collected from a cohort of 51 elderly subjects recruited from the Memory Clinic at Roozbeh Hospital in Tehran, Iran. DTI data were acquired from 9 white matter tracts, including the left and right arcuate fasciculus, left and right cingulum (frontal-parietal portion), left and right superior longitudinal fasciculus, and the corpus callosum (genu, body, and splenium). Relevant demographic, clinical, lifestyle, and medication data was gathered through in-person interviews conducted by trained researchers. Investigators received careful training in data collection protocols to maximize accuracy and consistency. The set of feature set for the input variables includes 49 variables from LIBRA and BALI. LIBRA comprised 42 variables in 4 categories: 1) socioeconomic demographic components, 2) components of physical health, 3) lifestyle components, and 4) laboratory components. (see

Appendix A). BALI provided 7 measurements: Gray matter lesions, subcortical dilated perivascular spaces, deep white matter lesions, periventricular white matter lesions, lesions in the basal ganglia and surrounding areas, lesions in the infratentorial regions, and global atrophy. The BALI factors denoting morphological changes were acquired through magnetic resonance imaging (see Appendix B).

Inclusion criteria were age over 50 years, and diagnosis of dementia. Exclusion criteria included the presence of major psychiatric illness, substance abuse, neurological comorbidities affecting cognition (eg tumor, infection), or contraindications to MRI. All participants or their surrogates gave their informed consent before data collection.

The university institutional review board approved the study protocol (approval number IR.IAU.SRB.REC.1401.285). This study was approved by the Research Ethics Committees of Islamic Azad University's Science and Research Branch. All research was carried out in accordance with relevant guidelines and regulations. Informed consent was obtained from all participants and/ or their legal guardians. All experiments were performed in accordance with relevant guidelines and regulations.

Diffusion Tensor Imaging

Diffusion tensor imaging (DTI) is an advanced magnetic resonance imaging technique that enables the visualization and quantification of water diffusion in biological tissues. DTI leverages diffusion-weighted MRI data to model diffusion as a tensor, describing the 3D diffusion properties in each image voxel [57, 58]. This allows us to estimate the microstructural organization and integrity of white matter pathways in the brain. S is a function of the diffusion coefficient (D), b-value, and the gradient direction (g) Eq. (1) [59]:

$$\frac{S(\mathbf{g}, \mathbf{b})}{S_0} = e^{-\mathbf{b} \cdot \mathbf{g}^T \mathbf{D} \mathbf{g}}, \quad (1)$$

D describes the diffusion properties of water molecules in each voxel in the tissue Eq. (2) [60]:

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}, \quad (2)$$

the tensor contains the primary diffusivities along the main axes (D_{xx} , D_{yy} , D_{zz}) as well as the off-diagonal elements representing correlations. From D, we can derive common diffusion metrics such as fractional anisotropy (FA) [61, 62].

FA describes the directional dependence of diffusion Eq. (3):

$$FA = \sqrt{\frac{1}{2} \frac{(\lambda_1 - \lambda_2)^2 + (\lambda_1 - \lambda_3)^2 + (\lambda_2 - \lambda_3)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \quad (3)$$

Where λ_1 , λ_2 , λ_3 are the tensor eigenvalues.

We performed DTI in each patient to quantify the progression of the disease, which served as the target output variable.

The disease progression measure was calculated from DTI-derived fractional anisotropy (FA) values in the white matter tracts [63]. Lower FA values indicate greater white matter disruption and more severe dementia. This served as the target variable for the supervised learning problem.

DTI characterizes the microstructural organization and integrity of the white matter pathways in the brain by modeling the diffusion properties of water molecules. Key DTI measures, such as fractional anisotropy (FA), can indicate white matter abnormalities associated with neurological disorders such as dementia. Details of DTI acquisition and processing methods can be found in [57, 58].

Data Pre-processing

An important step was to standardize each clinical factor evaluated for each subject, including the LIBRA and BALI measures described in Appendices 1 and 2. Although necessary for analysis, a challenge is that these factors have widely varying scales and ranges [64].

To allow an impartial analysis, I standardized all clinical data in the common range of 0 to 1 using the following formula Eq. (4):

$$x_s = \frac{x - \min(X)}{\max(X) - \min(X)}, \quad (4)$$

where s is the standardized value, x is the original numeric value, and X is the set of all x values [65].

This standardization ensures that all factors and measures are on the same scale, while preserving the distribution and relationship between values. With all data standardized, I can impartially analyze the relationships between clinical evaluations and imaging measures.

In the data pre-processing section, standardizing the clinical factors assessed for every subject, including the LIBRA and BALI measures, is a crucial step prior to analysis. Standardization ensures that features with inherently different scales and units of measurement are transformed into a common range, between 0 and 1 [66]. This is essential because machine learning algorithms can be unduly influenced by features with larger numeric ranges, leading to biased results and poor generalization. By standardizing the data, we ensure that all features contribute equally to the analysis, regardless of their original scales [67, 68]. Consequently, the model can impartially learn patterns and relationships without being disproportionately influenced by features with larger magnitudes. Standardization preserves the relative distribution and relationships between data points while ensuring fair consideration of all predictors. This crucial preprocessing step establishes a robust foundation for unbiased analysis and reliable prediction of dementia severity from the diverse set of clinical and neuroimaging characteristics [69].

Recursive Feature Elimination (RFE)

RFE is a popular feature selection technique that removes the least important features recursively [70, 71]. It can be used with different base machine learning models that assign feature importance scores [72].

Linear regression can be one such base model for RFE. In linear regression, the feature coefficients reflect the importance of that feature in predicting the target variable [73].

The linear regression model is defined as Eq. (5):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon, \quad (5)$$

where y is the target variable, x_1 to x_p are the p predictor features, β_0 is the intercept, β_1 to β_p is the feature coefficients, and ε is the error term [74, 75].

The importance score for each feature x_j can then be calculated as the absolute value of its coefficient β_j Eq. (6):

$$x_j = |\beta_j|, \quad (6)$$

at each RFE iteration, the feature with the lowest importance score is removed. The model is reconstructed on the remaining features, and the process is repeated until the desired number of features remains [74, 75].

Therefore, linear regression RFE involves recursively fitting linear regression, calculating the importance of feature as coefficient magnitudes, removing the least important feature, and refitting the model until the optimal set of features is obtained. This provides a computationally efficient way to select predictive features for the target variable using linear regression as the base model within RFE [76].

RFE selected a subset of the 10 most influential factors in the measurement of disease progression derived from DTI in Figure 1.

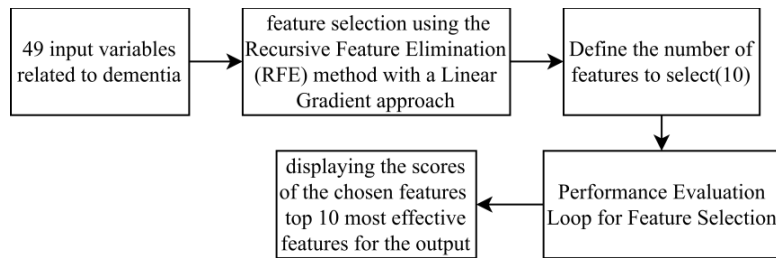


Figure 1. Block diagram of the proposed RFE algorithm

The RFE iteratively fits a linear regression model on the 49 BALI/ LIBRA input features, ranks the features by coefficient magnitude, and removes the least important feature. This process is repeated until the top 10 most predictive features for the DTI-derived disease progression measure are selected in Figure 1.

RFE is a powerful feature selection technique that iteratively removes irrelevant or redundant features while retaining the most predictive subset. This reduces dimensionality, improves model performance, and improves interpretability.

Input Features and Feature Selection

The input variables comprised 49 features from the BALI (7 factors) and LIBRA (42 factors) assessments (see Appendices A and B). RFE was applied to identify the top 10 most predictive features of this set of 49 variables to estimate the DTI-derived disease progression measure. RFE iteratively fits a linear regression model, ranks features by coefficient magnitude, and removes the least important feature until the desired number remains. The choice of 10 features strikes a balance between preserving informative predictors and reducing dimensionality.

Decision Tree

A decision tree is a powerful machine learning algorithm used for both prediction and regression tasks [77]. It is a tree-like structure where each internal node represents a feature (or attribute), each branch represents a decision based on that feature, and each leaf node represents an outcome or prediction [78]. The top node of the tree is called the root node, and represents the entire dataset [78]. To create branches, the data set is split into subsets based on a feature. The chosen feature is the one that provides the best data separation. This process is repeated recursively for

each subset, creating a tree structure [79]. At each internal node, a decision is made on the basis of a feature's value. The tree continues to branch until it reaches the leaf nodes. Entropy measures the impurity of a dataset, and information gain quantifies the reduction in impurity by splitting. Decision trees can become overly complex, which may lead to overfitting [80]. Pruning is a process in which some branches are removed to make the tree simpler and more generalizable. One of the advantages of decision trees is their interpretability. You can easily trace the path from the root to a leaf to understand how a decision was reached [79].

The decision tree model itself was created using the 'fitrtree' function [81]. This algorithm essentially learns from the training data, building a tree-like structure that maps the input features to the desired output. It is a visual representation of decision rules that can help predict outcomes based on new, unseen data. To evaluate the performance of the model, predictions were made on the validation and test sets [78]. The error was calculated by measuring the difference between the predicted and actual values. The error calculation was kept simple, using the mean absolute error. These errors are essential indicators of effectiveness and reliability [82].

The governing equations and explanations for decision tree creation using the fitrtree function in MATLAB:

Entropy (or Gini impurity) Calculation:

The decision tree algorithm calculates entropy ($H(S)$) or Gini impurity for the target variable (Y_{Train}) based on the class label distribution. Entropy is given by Eq. (7) [83]:

$$H(S) = -\sum (p_i \log_2(p_i)) , \quad (7)$$

where p_i represents the proportion of instances belonging to class i . The gini impurity follows a similar concept [83].

For each possible split point in the predictor variables (X_{Train}), the algorithm computes the information gain (IG) or Gini gain. These metrics quantify the reduction in entropy or Gini impurity, showing how much the split improves data purity Eq. (8) [78, 84]:

$$IG = H(S) - \sum (|S_v|/|S| (H(S_v))) , \quad (8)$$

where S is the current dataset, (S_v) are the subsets created by the split, and $H(S_v)$ is the entropy or Gini impurity of subset, (S_v) [85, 86].

The algorithm selects the split point that maximizes the information gain (or the Gini gain) to partition the data at each node in the tree. Evaluates all potential splits in the predictor variables.

The process is applied recursively to construct a tree structure. The internal nodes represent the split points based on the predictor variables, and the leaf nodes represent values for predictions [87].

We trained a decision tree model to predict the severity of Alzheimer's based on DTI. At each split, the decision tree optimized information gain, measuring the reduction in entropy after dividing the data. This process sculpted the complex data into an interpretable tree structure in Figure 2.

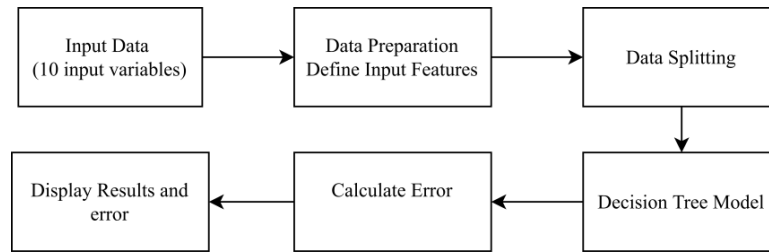


Figure 2. Block diagram of the proposed decision tree algorithm

Block diagram of the proposed decision tree algorithm. The decision tree recursively splits the data based on the 10 selected BALI/ LIBRA features to maximize information gain (reduction in entropy/ impurity). The final tree structure maps the input features to the predictions of the DTI-derived disease progression measure in Figure 2.

Decision trees are machine learning models that recursively split data based on input features to maximize the separation of classes or prediction accuracy. Their hierarchical tree structure and intuitive rules make them highly interpretable.

The proposed machine learning framework integrates multimodal data through an innovative two-stage algorithm. The first stage applies Recursive Feature Elimination (RFE) to derive a compact set of top predictive features from the high-dimensional BALI and LIBRA factors [88]. RFE iteratively fits a linear regression model, ranking all features by the magnitude of their coefficients [89]. At each iteration, the least important feature is eliminated until a predetermined number of surviving features remain [90]. This data-driven feature selection approach identifies the most informative lifestyle and structural factors while reducing dimensionality. In the second stage, the selected top features are used to train an interpretable decision tree model for dementia prediction. Decision trees recursively partition the data space by maximizing the information gain at each node, sculpting the complex multivariate distributions into a hierarchical tree structure [91]. This integrative approach synergistically combines complementary data modalities while enhancing interpretability through the decision tree framework.

Results

To identify the most predictive factors, we applied RFE on the 49 variables, which iteratively prunes the least important features. RFE selected a subset of the 10 most influential factors in the measurement of disease progression derived from DTI. These top-ranking biomarkers were used to develop the final decision tree model. The list of the top 10 factors of the input variables of LIBRA and BALI relative to each of the DTI tracts as output with importance scores is shown in Table 1.

Table 1 lists the top 10 BALI and LIBRA factors selected by RFE, ranked by their importance scores (regression coefficients) to predict in the FA values for each DTI tract.

Table 1. List of top 10 features with importance scores using RFE

Diffusion Tensor Imaging (FA)											
Arcuate_Fasciculus_L	Factors	6	4	38	15	8	24	46	44	21	20
	RFE	0.178754	0.252192	0.409541	0.448106	0.472541	0.508274	0.524891	0.543831	0.577149	0.587064
Arcuate_Fasciculus_R	Factors	35	27	9	22	47	45	24	4	10	38
	RFE	0.091903	0.149444	0.195241	0.257047	0.305354	0.346572	0.380917	0.431152	0.445082	0.4586
Cingulum_Frontal_Parietal_L	Factors	32	1	35	10	18	21	6	27	20	45
	RFE	0.166296	0.283323	0.363448	0.419051	0.503561	0.529305	0.552518	0.588751	0.635156	0.654975
Cingulum_Frontal_Parietal_R	Factors	6	17	2	32	30	18	39	35	27	25
	RFE	0.192893	0.384751	0.506899	0.556467	0.593786	0.611621	0.633986	0.649885	0.663591	0.676541
Superior_Longitudinal_Fasciculus_L	Factors	49	8	20	38	24	4	16	34	17	6
	RFE	0.18475	0.341002	0.391806	0.510035	0.565731	0.592481	0.613366	0.638113	0.667556	0.687449
Superior_Longitudinal_Fasciculus_R	Factors	2	38	24	20	8	45	39	40	9	18
	RFE	0.110333	0.23742	0.290458	0.335593	0.38878	0.4169	0.446979	0.458735	0.483564	0.526755
Corpus_Callosum_GEN	Factors	6	29	9	31	1	39	48	20	35	16
	RFE	0.22969	0.314046	0.354564	0.378597	0.393811	0.422378	0.441972	0.455895	0.470908	0.521148
Corpus_Callosum_Body	Factors	6	13	27	22	20	35	8	17	42	28
	RFE	0.258628	0.35093	0.455137	0.48703	0.521242	0.588952	0.603625	0.626639	0.637887	0.646875
Corpus_Callosum_SPL	Factors	6	29	20	3	35	16	7	21	18	27
	RFE	0.139123	0.216037	0.294052	0.335871	0.397728	0.469177	0.498992	0.530716	0.54147	0.552176

Using these 10 LIBRA and BALI factors as predictors, we trained a decision tree model to predict the severity of Alzheimer's based on DTI. The resulting decision tree model demonstrated high precision in predicting the level of disease derived from DTI from the selected factors. The decision tree model achieved 96.25% precision in detecting dementia in the independent test set. These criteria determine the model's ability to accurately predict disease stage based on biomarkers and brain imaging, and demonstrate potential clinical utility for predicting DTI and ultimately diagnosing dementia.

The recursive feature elimination strategy employed in this study elucidates intriguing insights into the key predictors of white matter microstructural integrity from the multidimensional BALI and LIBRA assessments. In particular, the global atrophy factor of the BALI index emerged as one of the top-ranking predictors in multiple DTI tracts, including the arcuate fasciculus, cingulum, superior longitudinal fasciculus, and corpus callosum (Table 1). Interestingly, the alcohol intake factor from LIBRA also exhibited consistent predictive utility for several DTI measures, suggesting a potential association between moderate alcohol consumption and white matter health. Conversely, the antipsychotic drugs factor from LIBRA emerged as a robust negative predictor, aligning with previous literature implicating certain antipsychotics in white matter abnormalities. These results provide novel insights into the intricate interplay between lifestyle factors, vascular pathology, and white matter integrity in the context of dementia, offering potential avenues for preventive interventions and personalized risk stratification. Smoking factor negatively predicted integrity in the corpus callosum, corroborating previous evidence linking smoking with disrupted callosal connectivity and interhemispheric communication.

Experimental Setup and Evaluation

The data set was randomly divided into training (70%), validation (15%), and test (15%) sets. A decision tree model was trained on the training set using the 10 selected BALI/ LIBRA features as predictors and the DTI-derived disease progression measure as target. The model hyperparameters were tuned in the validation set to prevent overfitting. Finally, the tuned model's performance was evaluated on the held-out test set to obtain an unbiased estimate of its generalization ability. The 96.25% accuracy on the test set indicates that the decision tree model can reliably predict dementia severity from the selected BALI/ LIBRA factors, without requiring costly acquisition of DTI. This enables cost-effective screening and monitoring of dementia progression.

Discussion

The ability to accurately stage the severity of dementia using machine learning could have important clinical implications. Objective quantification of progression could help guide prognosis and counseling for patients and their families. Predicted severity could also inform treatment decisions and the timing of interventions. For mild dementia, the model could justify the introduction of medications to manage symptoms and prevent decline. For severe dementia, clinical utility includes planning for safety, caregiver needs, and transition to long-term care. Furthermore, the model could identify rapid progressors for more aggressive intervention and monitoring.

There are several opportunities to improve the performance and utility of the model. Expanding to larger and more diverse cohorts of patients will improve robustness and generalizability. Adding modalities such as resting-state fMRI and FDG-PET could incorporate functional and metabolic biomarkers. Ultimately, the integrative modeling framework demonstrated here could be extended to user-friendly tools for neurologists to use biomarkers in early and accurate diagnosis.

A key limitation of this study is the relatively small sample size of 51 participants. While the results are promising, larger multicenter studies are needed to validate the generalizability of the proposed approach across diverse populations. Furthermore, the current model only considers cross-sectional data; incorporating longitudinal measurements could potentially improve predictions of disease progression over time.

Our integrative machine learning framework is based on previous studies that have explored BALI, LIBRA, and DTI individually for the assessment of dementia. However, most previous work has focused on single modalities, lacking the complementary information gained by combining structural, microstructural, and lifestyle factors. Multimodal data fusion using machine learning has shown promise in recent studies of Alzheimer's disease. Compared to existing methods, our approach is novel in strategically combining inexpensive and routinely collected BALI/ LIBRA factors to estimate costly DTI metrics, enabling scalable screening and monitoring applications. With further validation, this cost-effective strategy could potentially be translated into clinical practice and large-scale studies that seek affordable biomarkers of dementia progression.

The present study is of profound significance for the neuroimaging and dementia research communities by pioneering a novel and cost-effective approach to assess the microstructural integrity of white matter, a crucial biomarker of dementia progression. Using readily available structural MRI and lifestyle data through advanced machine learning techniques, our framework circumvents the need for widespread acquisitions of DTI, which are often impeded by practical and logistical constraints. This accessibility paves the way for larger-scale studies and routine screening paradigms, facilitating early detection and monitoring of white matter abnormalities associated with dementia. Furthermore, the elucidation of specific BALI and LIBRA factors as robust predictors of regional white matter integrity provides invaluable information on the complex interplay between vascular pathology, lifestyle factors, and neurodegeneration. These findings not only advance our mechanistic understanding of the etiology of dementia but also underscore potential targets for preventive interventions and lifestyle modifications, aligning with the growing emphasis on personalized and precision medicine approaches in neurodegenerative disorders.

Furthermore, the interpretable nature of the decision tree model used in this study offers a unique advantage, enabling researchers and clinicians to understand the intricate decision rules and multivariate relationships linking structural, lifestyle, and microstructural markers to the severity of dementia. This transparency facilitates the translation of complex multimodal neuroimaging data into actionable insights, informing individualized risk stratification, treatment planning, and patient counseling. Collectively, our integrative framework paves the way for future studies investigating the dynamic interaction between modifiable risk factors, brain pathology, and white matter degeneration, ultimately contributing to the development of targeted personalized interventions for the prevention and treatment of dementia.

Conclusions

This study demonstrates the proof-of-concept for an integrative framework that combines multimodal neuroimaging biomarkers and machine learning to predict the severity of dementia. Our approach leveraged the complementary strengths of BALI, LIBRA, and DTI to comprehensively characterize the neuropathology of dementia. The application of recursive feature elimination and decision trees enabled predictive modeling from high-dimensional imaging data. This sets the stage for larger-scale validation and potential clinical.

This study provides initial evidence that the combination of neuroimaging modalities and machine learning could eventually enable earlier and more accurate dementia diagnosis in clinical practice. Following further validation in larger, independent cohorts, our predictive modeling framework could help identify high-risk individuals in the earliest stages when interventions are most effective. By detecting progression years before obvious symptoms emerge, neuroimaging biomarkers and informatics offer hope for early treatment and prevention when dementia remains reversible.

In summary, our work pioneers a path toward precision approaches for early dementia prediction using neuroimaging biomarkers and advanced analytics. Much progress remains to be made through rigorous science and partnerships. With collective bold vision and dedication, transformative solutions to this global health crisis are attainable. Tapping the tremendous advances in biotechnology and informatics gives real hope to unravel the complexity of dementia. This study provides an important starting point for the dementia research community to move forward together. It will take a large collaborative effort across disciplines to fully leverage integrative diagnostics for early interventions. We must continue pushing boundaries without forgetting the lives these technologies aim to improve. With compassion and commitment to open science, a brighter future for dementia care is within reach.

Data Availability Statement

Data sets generated during the current study are available from the corresponding author on reasonable request.

References

- [1] E. Brai, A. Tonacci, V. Brugada-Ramentol, F. D'Andrea, and L. Alberi, "Intercepting dementia: Awareness and innovation as key tools," *Frontiers in Aging Neuroscience*, vol. 13, p. 730727, 2021.
- [2] J. Ryan, P. Fransquet, J. Wrigglesworth, and P. Lacaze, "Phenotypic heterogeneity in dementia: a challenge for epidemiology and biomarker studies," *Frontiers in public health*, vol. 6, p. 181, 2018.
- [3] E. Nichols *et al.*, "Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019," *The Lancet Public Health*, vol. 7, no. 2, pp. e105-e125, 2022.
- [4] A. Nandi *et al.*, "Global and regional projections of the economic burden of Alzheimer's disease and related dementias from 2019 to 2050: A value of statistical life approach," *EClinicalMedicine*, vol. 51, 2022.
- [5] A. Wimo *et al.*, "The worldwide costs of dementia in 2019," *Alzheimer's & Dementia*, 2023.
- [6] M. P. Aranda *et al.*, "Impact of dementia: Health disparities, population trends, care interventions, and economic costs," *Journal of the American Geriatrics Society*, vol. 69, no. 7, pp. 1774-1783, 2021.
- [7] D. W. Shineman and H. M. Fillit, "Novel strategies for the prevention of dementia from Alzheimer's disease," *Dialogues in clinical neuroscience*, 2022.
- [8] A. Porsteinsson, R. Isaacson, S. Knox, M. Sabbagh, and I. Rubino, "Diagnosis of early Alzheimer's disease: clinical practice in 2021," *The journal of prevention of Alzheimer's disease*, vol. 8, pp. 371-386, 2021.

- [9] P. D. Harvey, "Clinical applications of neuropsychological assessment," *Dialogues in clinical neuroscience*, 2022.
- [10] W. M. van Oostveen and E. C. de Lange, "Imaging techniques in Alzheimer's disease: a review of applications in early diagnosis and longitudinal monitoring," *International journal of molecular sciences*, vol. 22, no. 4, p. 2110, 2021.
- [11] F. Baldacci *et al.*, "Blood-based biomarker screening with agnostic biological definitions for an accurate diagnosis within the dimensional spectrum of neurodegenerative diseases," *Biomarkers for Alzheimer's Disease Drug Development*, pp. 139-155, 2018.
- [12] H. Jiang, Z. Lu, M. B. Gerdroodbary, A. Sabernaemi, and S. Salavatidezfouli, "The influence of sac centreline on saccular aneurysm rupture: Computational study," *Scientific Reports*, vol. 13, no. 1, p. 11288, 2023.
- [13] A. Sabernaemi, M. Barzegar Gerdroodbary, S. Salavatidezfouli, and P. Valipour, "Influence of stent-induced vessel deformation on hemodynamic feature of bloodstream inside ICA aneurysms," *Biomechanics and Modeling in Mechanobiology*, vol. 22, no. 4, pp. 1193-1207, 2023.
- [14] B. Alicioglu and H. M. Karakas, "Brain aging: radiological biomarkers," in *Assessments, Treatments and Modeling in Aging and Neurological Disease*: Elsevier, 2021, pp. 235-247.
- [15] M. Modat, D. M. Cash, L. Dos Santos Canas, M. Bocchetta, and S. Ourselin, "Machine Learning for Alzheimer's Disease and Related Dementias," in *Machine Learning for Brain Disorders*: Springer, 2023, pp. 807-846.
- [16] Y. Zhou, Z. Song, X. Han, H. Li, and X. Tang, "Prediction of Alzheimer's disease progression based on magnetic resonance imaging," *ACS Chemical Neuroscience*, vol. 12, no. 22, pp. 4209-4223, 2021.
- [17] L. A. Grajauskas, W. Siu, G. Medvedev, H. Guo, R. C. D'Arcy, and X. Song, "MRI-based evaluation of structural degeneration in the ageing brain: Pathophysiology and assessment," *Ageing research reviews*, vol. 49, pp. 67-82, 2019.
- [18] C. D. Mayo *et al.*, "Relationship between DTI metrics and cognitive function in Alzheimer's disease," *Frontiers in aging neuroscience*, vol. 10, p. 436, 2019.
- [19] M. C. Power *et al.*, "Association of white matter microstructural integrity with cognition and dementia," *Neurobiology of aging*, vol. 83, pp. 63-72, 2019.
- [20] D. Xiao, K. Wang, L. Theriault, E. Charbel, and A. s. D. N. Initiative, "White matter integrity and key structures affected in Alzheimer's disease characterized by diffusion tensor imaging," *European Journal of Neuroscience*, vol. 56, no. 8, pp. 5319-5331, 2022.
- [21] R. Laforce Jr *et al.*, "Molecular imaging in dementia: Past, present, and future," *Alzheimer's & Dementia*, vol. 14, no. 11, pp. 1522-1552, 2018.
- [22] J.-C. Park *et al.*, "A logical network-based drug-screening platform for Alzheimer's disease representing pathological features of human brain organoids," *Nature Communications*, vol. 12, no. 1, p. 280, 2021.
- [23] P. Yang, G. Bi, J. Qi, X. Wang, Y. Yang, and L. Xu, "Multimodal wearable intelligence for dementia care in healthcare 4.0: A survey," *Information Systems Frontiers*, pp. 1-18, 2021.
- [24] L. J. Dominguez *et al.*, "Nutrition, physical activity, and other lifestyle factors in the prevention of cognitive decline and dementia," *Nutrients*, vol. 13, no. 11, p. 4080, 2021.
- [25] J. V. Woodside *et al.*, "Meeting report: plant-rich dietary patterns and health," *Proceedings of the Nutrition Society*, vol. 81, no. 4, pp. 288-305, 2022.
- [26] M. Brasure *et al.*, "Physical activity interventions in preventing cognitive decline and Alzheimer-type dementia: a systematic review," *Annals of internal medicine*, vol. 168, no. 1, pp. 30-38, 2018.
- [27] S. B. Félix, O. Ribeiro, and H. Maia, "Personalized cognitive stimulation through personhood: A case report on dementia diagnosis acceptance and therapeutic engagement," *Clinical Gerontologist*, vol. 43, no. 2, pp. 233-239, 2020.
- [28] V. Hoel, C. M. Feunou, and K. Wolf-Ostermann, "Technology-driven solutions to prompt conversation, aid communication and support interaction for people with dementia and their caregivers: a systematic literature review," *BMC geriatrics*, vol. 21, pp. 1-11, 2021.
- [29] K. K. Gulia and V. M. Kumar, "Sleep disorders in the elderly: a growing challenge," *Psychogeriatrics*, vol. 18, no. 3, pp. 155-165, 2018.
- [30] Y. Xu *et al.*, "Exploring patient medication adherence and data mining methods in clinical big data: A contemporary review," *Journal of Evidence-Based Medicine*, 2023.
- [31] N. Falgàs *et al.*, "Hippocampal atrophy has limited usefulness as a diagnostic biomarker on the early onset Alzheimer's disease patients: A comparison between visual and quantitative assessment," *NeuroImage: Clinical*, vol. 23, p. 101927, 2019.

- [32] J. F. Hodes *et al.*, "Alzheimer's "prevention" vs."risk reduction": transcending semantics for clinical practice," *Frontiers in neurology*, vol. 9, p. 1179, 2019.
- [33] M. Bruun *et al.*, "Detecting frontotemporal dementia syndromes using MRI biomarkers," *NeuroImage: Clinical*, vol. 22, p. 101711, 2019.
- [34] D. Aarsland, "Epidemiology and pathophysiology of dementia-related psychosis," *The Journal of clinical psychiatry*, vol. 81, no. 5, p. 27625, 2020.
- [35] S. Moguilner *et al.*, "Multi-feature computational framework for combined signatures of dementia in underrepresented settings," *Journal of Neural Engineering*, vol. 19, no. 4, p. 046048, 2022.
- [36] C. Fabrizio, A. Termine, C. Caltagirone, and G. Sancesario, "Artificial intelligence for Alzheimer's disease: promise or challenge?," *Diagnostics*, vol. 11, no. 8, p. 1473, 2021.
- [37] M. Zitnik, F. Nguyen, B. Wang, J. Leskovec, A. Goldenberg, and M. M. Hoffman, "Machine learning for integrating data in biology and medicine: Principles, practice, and opportunities," *Information Fusion*, vol. 50, pp. 71-91, 2019.
- [38] B. Bratić, V. Kurbalija, M. Ivanović, I. Oder, and Z. Bosnić, "Machine learning for predicting cognitive diseases: methods, data sources and risk factors," *Journal of medical systems*, vol. 42, pp. 1-15, 2018.
- [39] M. Niazi and G. Niedbala, "Machine learning for plant breeding and biotechnology," *Agriculture*, vol. 10, no. 10, p. 436, 2020.
- [40] M. A. Myszczyńska *et al.*, "Applications of machine learning to diagnosis and treatment of neurodegenerative diseases," *Nature Reviews Neurology*, vol. 16, no. 8, pp. 440-456, 2020.
- [41] A. Javeed, A. L. Dallora, J. S. Berglund, and P. Anderberg, "An intelligent learning system for unbiased prediction of dementia based on autoencoder and adaboost ensemble learning," *Life*, vol. 12, no. 7, p. 1097, 2022.
- [42] M. Mohammadzaheri *et al.*, "Modelling of engineering systems with small data: A comparative study," in *Perspectives and Considerations on the Evolution of Smart Systems*: IGI Global, 2023, pp. 120-136.
- [43] M. Mohammadzaheri *et al.*, "Adaptive Charge Estimation of Piezoelectric Actuators with a Variable Sensing Resistor, an Artificial Intelligence Approach," *Engineering Letters*, vol. 30, no. 1, 2022.
- [44] M. Bucholtz *et al.*, "Artificial intelligence for dementia research methods optimization," *Alzheimer's & Dementia*, 2023.
- [45] S. Grueso and R. Viejo-Sobera, "Machine learning methods for predicting progression from mild cognitive impairment to Alzheimer's disease dementia: a systematic review," *Alzheimer's research & therapy*, vol. 13, pp. 1-29, 2021.
- [46] D. P. Veitch *et al.*, "Using the Alzheimer's Disease Neuroimaging Initiative to improve early detection, diagnosis, and treatment of Alzheimer's disease," *Alzheimer's & Dementia*, vol. 18, no. 4, pp. 824-857, 2022.
- [47] M. Jensen, T. Zeller, R. Twerenbold, and G. Thomalla, "Circulating cardiac biomarkers, structural brain changes, and dementia: Emerging insights and perspectives," *Alzheimer's & Dementia*, vol. 19, no. 4, pp. 1529-1548, 2023.
- [48] A. Akbarifar, A. Maghsoudpour, F. Mohammadian, M. Mohammadzaheri, and O. Ghaemi, "Fuzzy Clustering to Assess BALI and LIBRA factors for Estimation of DTI measures," in *2023 28th International Conference on Automation and Computing (ICAC)*, 2023, pp. 1-6: IEEE.
- [49] A. Akbarifar, A. Maghsoudpour, F. Mohammadian, M. Mohammadzaheri, and O. Ghaemi, "Predicting dementia progression with fully connected cascade neural networks," 2024.
- [50] S. Ballav, B. Biswas, S. Dey, V. K. Sahu, and S. Basu, "Image processing: an early detection tool for Alzheimer's disease," in *Smart Diagnostics for Neurodegenerative Disorders*: Elsevier, 2024, pp. 99-136.
- [51] J. Arantes *et al.*, "Dissolving the dichotomies between online and campus-based teaching: a collective response to The Manifesto for Teaching Online," *Postdigital Science and Education*, 2021.
- [52] G. Castellazzi *et al.*, "A machine learning approach for the differential diagnosis of Alzheimer and vascular dementia fed by MRI selected features," *Frontiers in neuroinformatics*, p. 25, 2020.
- [53] N. S. Graham and D. J. Sharp, "Understanding neurodegeneration after traumatic brain injury: from mechanisms to clinical trials in dementia," *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 90, no. 11, pp. 1221-1233, 2019.
- [54] D. Beck *et al.*, "White matter microstructure across the adult lifespan: A mixed longitudinal and cross-sectional study using advanced diffusion models and brain-age prediction," *NeuroImage*, vol. 224, p. 117441, 2021.
- [55] J. M. Groot, "Neural mechanisms of the wandering mind," 2023.

- [56] M. M. Alsaleh *et al.*, "Prediction of disease comorbidity using explainable artificial intelligence and machine learning techniques: A systematic review," *International Journal of Medical Informatics*, p. 105088, 2023.
- [57] J. Huang, L. Wang, C. Chu, W. Liu, and Y. Zhu, "Accelerating cardiac diffusion tensor imaging combining local low-rank and 3D TV constraint," *Magnetic Resonance Materials in Physics, Biology and Medicine*, vol. 32, pp. 407-422, 2019.
- [58] Q. Tian *et al.*, "DeepDTI: High-fidelity six-direction diffusion tensor imaging using deep learning," *NeuroImage*, vol. 219, p. 117017, 2020.
- [59] D. C. Karampinos, S. Banerjee, K. F. King, T. M. Link, and S. Majumdar, "Considerations in high-resolution skeletal muscle diffusion tensor imaging using single-shot echo planar imaging with stimulated-echo preparation and sensitivity encoding," *NMR in Biomedicine*, vol. 25, no. 5, pp. 766-778, 2012.
- [60] V. Mazzoli, K. Moulin, F. Kogan, B. A. Hargreaves, and G. E. Gold, "Diffusion tensor imaging of skeletal muscle contraction using oscillating gradient spin Echo," *Frontiers in Neurology*, vol. 12, p. 608549, 2021.
- [61] J. Guerrero, T. A. Gallagher, A. L. Alexander, and A. S. Field, "Diffusion Tensor Magnetic Resonance Imaging—Physical Principles," in *Functional Neuroradiology: Principles and Clinical Applications*: Springer, 2023, pp. 903-932.
- [62] A. Akbarifar, A. Maghsoudpour, F. Mohammadian, M. Mohammadzaheri, and O. Ghaemi, "Multimodal Dementia Identification using Lifestyle and Brain Lesions, a Machine Learning Approach," *AIP Advances*, vol. 14, no. 6, 2024.
- [63] A. Akbarifar, A. Maghsoudpour, F. Mohammadian, M. Mohammadzaheri, and O. Ghaemi, "A Novel Approach to Dementia Prediction Leveraging Recursive Feature Elimination and Decision Tree," 2024.
- [64] R. A. Maronna, R. D. Martin, V. J. Yohai, and M. Salibián-Barrera, *Robust statistics: theory and methods (with R)*. John Wiley & Sons, 2019.
- [65] M. Mazziotta and A. Pareto, "Normalization methods for spatio-temporal analysis of environmental performance: Revisiting the Min–Max method," *Environmetrics*, vol. 33, no. 5, p. e2730, 2022.
- [66] I. E. STANDARDIZED, "INTERPRETATION OF STANDARDIZED PERCENTILES," *A Compendium of Neuropsychological Tests: Fundamentals of Neuropsychological Assessment and Test Reviews for Clinical Practice*, p. 2, 2023.
- [67] A. Akbarifar *et al.*, "Fault Detection of Gas Pipelines Using Mechanical Waves and Intelligent Techniques."
- [68] M. M. Mohammadzaheri, A. Akbarifar, and M. Ghodsi, "Health monitoring of welded pipelines with mechanical waves and fuzzy inference systems," in *16th International Gas Union Research Conference*, 2020: International Gas Union.
- [69] L. F. Bachman, *Statistical analyses for language assessment book*. Cambridge University Press, 2004.
- [70] G. Karami, M. Giuseppe Orlando, A. Delli Pizzi, M. Caulo, and C. Del Gratta, "Predicting overall survival time in glioblastoma patients using gradient boosting machines algorithm and recursive feature elimination technique," *Cancers*, vol. 13, no. 19, p. 4976, 2021.
- [71] W. Liu and J. Wang, "Recursive elimination–election algorithms for wrapper feature selection," *Applied Soft Computing*, vol. 113, p. 107956, 2021.
- [72] L. Li, W.-K. Ching, and Z.-P. Liu, "Robust biomarker screening from gene expression data by stable machine learning-recursive feature elimination methods," *Computational biology and chemistry*, vol. 100, p. 107747, 2022.
- [73] X. Liu, H. Tang, Y. Ding, and D. Yan, "Investigating the performance of machine learning models combined with different feature selection methods to estimate the energy consumption of buildings," *Energy and Buildings*, vol. 273, p. 112408, 2022.
- [74] M. Lee, J.-H. Lee, and D.-H. Kim, "Gender recognition using optimal gait feature based on recursive feature elimination in normal walking," *Expert Systems with Applications*, vol. 189, p. 116040, 2022.
- [75] W. Lian, G. Nie, B. Jia, D. Shi, Q. Fan, and Y. Liang, "An intrusion detection method based on decision tree-recursive feature elimination in ensemble learning," *Mathematical Problems in Engineering*, vol. 2020, pp. 1-15, 2020.
- [76] E. M. Senan *et al.*, "Diagnosis of chronic kidney disease using effective classification algorithms and recursive feature elimination techniques," *Journal of Healthcare Engineering*, vol. 2021, 2021.
- [77] I. D. Mienye, Y. Sun, and Z. Wang, "Prediction performance of improved decision tree-based algorithms: a review," *Procedia Manufacturing*, vol. 35, pp. 698-703, 2019.

- [78] T. Thomas, A. P. Vijayaraghavan, S. Emmanuel, T. Thomas, A. P. Vijayaraghavan, and S. Emmanuel, "Applications of decision trees," *Machine learning approaches in cyber security analytics*, pp. 157-184, 2020.
- [79] V. G. Costa and C. E. Pedreira, "Recent advances in decision trees: An updated survey," *Artificial Intelligence Review*, vol. 56, no. 5, pp. 4765-4800, 2023.
- [80] Y. Wang, D. Wang, N. Geng, Y. Wang, Y. Yin, and Y. Jin, "Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection," *Applied Soft Computing*, vol. 77, pp. 188-204, 2019.
- [81] J. Yang, C. Xiang, and J. Liu, "Clinical significance of combining salivary mRNAs and carcinoembryonic antigen for ovarian cancer detection," *Scandinavian Journal of Clinical and Laboratory Investigation*, vol. 81, no. 1, pp. 39-45, 2021.
- [82] A. Ahmad *et al.*, "Prediction of compressive strength of fly ash based concrete using individual and ensemble algorithm," *Materials*, vol. 14, no. 4, p. 794, 2021.
- [83] B. Jiang, M. Durieux, and K. V. Golubnichiy, "Solving the Stock Option Forecast problem by a numerical method for the Black-Scholes Equation with Machine Learning Classification Model," *arXiv preprint arXiv:2209.03512*, 2022.
- [84] A. K. Martinsen, "Predicting Treatment Outcome Using Interpretable Models for Patients with Head and Neck Cancer," Norwegian University of Life Sciences, 2023.
- [85] V. Kushwaha, "Breast cancer diagnostic using machine learning: applying supervised learning techniques to Coimbra and Wisconsin datasets," 2023.
- [86] H. Ratnavath and V. Narasimha, "NETWORK INTRUSION DETECTION USING ENSEMBLE WEIGHTED VOTING CLASSIFIER BASED HONEY POT AND IDS FRAMEWORK," *NETWORK*, vol. 52, no. 3, 2023.
- [87] I. Podhorska, J. Vrbka, G. Lazaroiu, and M. Kovacova, "Innovations in financial management: Recursive prediction model based on decision trees," 2020.
- [88] Y. Wang and Y. Li, "Mapping the ratoon rice suitability region in China using random forest and recursive feature elimination modeling," *Field Crops Research*, vol. 301, p. 109016, 2023.
- [89] F. G. F. Niquini *et al.*, "Recursive Feature Elimination and Neural Networks Applied to the Forecast of Mass and Metallurgical Recoveries in A Brazilian Phosphate Mine," *Minerals*, vol. 13, no. 6, p. 748, 2023.
- [90] Y. Tang, Y.-Q. Zhang, and Z. Huang, "Development of two-stage SVM-RFE gene selection strategy for microarray expression data analysis," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 4, no. 3, pp. 365-381, 2007.
- [91] J. Zhang, X. Wang, U. Kruger, and F.-Y. Wang, "Principal curve algorithms for partitioning high-dimensional data spaces," *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 367-380, 2010.