



OPEN ACCESS

EDITED BY

Man Fai Leung,
Anglia Ruskin University, United Kingdom

REVIEWED BY

Nabil Abdulhafiz Alhakamy,
King Abdulaziz University, Saudi Arabia
Saifullah Tumrani,
Heidelberg University, Germany
Shahzad Akbar,
Riphah International University, Pakistan

*CORRESPONDENCE

Johan Sanmartin Berglund,
✉ johan.sanmartin.berglund@bth.se

RECEIVED 10 November 2023

ACCEPTED 05 December 2023

PUBLISHED 08 January 2024

CITATION

Javeed A, Anderberg P, Ghazi AN, Noor A,
Elmståhl S and Berglund JS (2024),
Breaking barriers: a statistical and
machine learning-based hybrid system
for predicting dementia.
Front. Bioeng. Biotechnol. 11:1336255.
doi: 10.3389/fbioe.2023.1336255

COPYRIGHT

© 2024 Javeed, Anderberg, Ghazi, Noor,
Elmståhl and Berglund. This is an open-
access article distributed under the terms
of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is
permitted, provided the original author(s)
and the copyright owner(s) are credited
and that the original publication in this
journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Breaking barriers: a statistical and machine learning-based hybrid system for predicting dementia

Ashir Javeed¹, Peter Anderberg^{1,2}, Ahmad Nauman Ghazi³,
Adeeb Noor⁴, Sölve Elmståhl⁵ and Johan Sanmartin Berglund^{1*}

¹Department of Health, Blekinge Institute of Technology, Karlskrona, Sweden, ²School of Health Sciences, University of Skövde, Skövde, Sweden, ³Department of Software Engineering, Blekinge Institute of Technology, Karlskrona, Sweden, ⁴Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia, ⁵EpiHealth: Epidemiology for Health, Lund University, SUS Malmö, Malmö, Sweden

Introduction: Dementia is a condition (a collection of related signs and symptoms) that causes a continuing deterioration in cognitive function, and millions of people are impacted by dementia every year as the world population continues to rise. Conventional approaches for determining dementia rely primarily on clinical examinations, analyzing medical records, and administering cognitive and neuropsychological testing. However, these methods are time-consuming and costly in terms of treatment. Therefore, this study aims to present a noninvasive method for the early prediction of dementia so that preventive steps should be taken to avoid dementia.

Methods: We developed a hybrid diagnostic system based on statistical and machine learning (ML) methods that used patient electronic health records to predict dementia. The dataset used for this study was obtained from the Swedish National Study on Aging and Care (SNAC), with a sample size of 43040 and 75 features. The newly constructed diagnostic extracts a subset of useful features from the dataset through a statistical method (F-score). For the classification, we developed an ensemble voting classifier based on five different ML models: decision tree (DT), naive Bayes (NB), logistic regression (LR), support vector machines (SVM), and random forest (RF). To address the problem of ML model overfitting, we used a cross-validation approach to evaluate the performance of the proposed diagnostic system. Various assessment measures, such as accuracy, sensitivity, specificity, receiver operating characteristic (ROC) curve, and Matthew's correlation coefficient (MCC), were used to thoroughly validate the devised diagnostic system's efficiency.

Results: According to the experimental results, the proposed diagnostic method achieved the best accuracy of 98.25%, as well as sensitivity of 97.44%, specificity of 95.744%, and MCC of 0.7535.

Discussion: The effectiveness of the proposed diagnostic approach is compared to various cutting-edge feature selection techniques and baseline ML models. From experimental results, it is evident that the proposed diagnostic system outperformed the prior feature selection strategies and baseline ML models regarding accuracy.

KEYWORDS

dementia, voting classifier, F-score, machine learning, feature selection

1 Introduction

Dementia is a severe neurological condition that causes memory, thinking, behavioral issues, and a steady decline in cognitive abilities (Creavin et al., 2016). Worldwide, millions of people are impacted by dementia. Around the globe today, 50 million people are thought to be affected by dementia. Dementia incidence is expected to triple by 2050. The prevalence of dementia will keep rising as the global population ages, putting enormous strain on healthcare systems worldwide (Iadecola, 2016). Early dementia diagnosis and prediction are crucial because they allow for quick intervention, improved patient treatment, and possible preventive measures; thus, prevention is essential for addressing this public health issue (Nichols et al., 2022).

The standard procedures for diagnosing dementia rely on clinical evaluations, which can be arbitrary and vulnerable to discrepancies between various assessors. These evaluations involve reviewing the medical records and conducting cognitive and neuropsychological tests Hsiu et al. (2022). Cognitive diagnostic tests or pathological characteristics diagnose dementia in its early stages. Pathological features can be found through neuroimaging. The alteration in neuronal structure is examined using magnetic resonance imaging (MRI) Studholme et al. (2004); Duchesne et al. (2008). These techniques are helpful but cannot address the minor alterations in brain activity that characterize dementia's early stages. Electroencephalography (EEG) is another method to assess individuals in the initial phases of dementia Ahiskali et al. (2009). EEG and MRI imaging were coupled by Patel et al. to enhance the identification of dementia in its early stages Patel et al. (2008). However, due to the unacceptably high cost of testing and the excessively drawn-out and intrusive nature of the testing process, such instruments are insufficient for diagnosing dementia. Additionally, new studies advise using computed tomography (CT) or MRI of the brain to rule out structural explanations for the clinical phenotype. According to estimates, primary care physicians misdiagnose between 29% and 76% of people with dementia or are likely to develop dementia Patnode et al. (2020). This highlights the critical need for new diagnostic methods to identify dementia in its earliest stages reliably. Recent years have seen the emergence of machine learning (ML) as a potent tool for predictive analytics and pattern identification, providing exciting chances for advancements in this field Tanveer et al. (2020). Large-scale data sets can be examined by ML algorithms, which can also reveal hidden patterns that were previously undetected and make predictions. Numerous data sources, such as genetic markers, brain imaging data, lifestyle factors, and neuropsychological tests, have been used by researchers to test the efficacy of ML for dementia detection (Basheer et al., 2021). There are substantial potential advantages to ML's capacity to identify dementia with early onset. Through prompt action when a disease is initially identified, it is possible to treat it optimally and enhance patients' quality of life. Early dementia prediction aids researchers in their search for novel biomarkers and drug targets that will enable them to create more potent treatments (Spooner et al., 2020). There are still difficulties utilizing ML for dementia prediction despite the encouraging findings. For the efficient training and validation of ML models, big, meticulously assembled datasets covering a variety of variables are required (Sivakani and Ansari, 2020).

The aim and purpose of this study is given as follows:

1. Constructing a dataset for dementia by integrating data from four distinct SNAC sites (Blekinge, Kungsholmen, and Skåne) employing data integration and harmonization criteria.
2. Significant features are selected from the dataset using a statistical method (F-score).
3. For the classification of dementia, an ensemble voting classifier based on DT, NB, SVM, LR, and RF was constructed.
4. The effectiveness of the proposed hybrid system, which combines an ensemble voting classifier and a statistical method (F-score), is also evaluated in comparison to three other feature selection techniques.
5. Experimental results show that the proposed model outperforms the baseline machine learning models, such as AdaBoost, Random Forest, Support Vector Machines, Linear Regression, Logistic Regression, Naive Bayes, and Decision Tree, according to the three commonly used evaluation metrics of accuracy, ROC curves, and AUC.

1.1 Literature review

Numerous studies have been conducted on applying ML approaches to solve problems across various medical applications and disease prediction. Researchers have developed several ML and deep learning (DL) based algorithms for the early prediction of dementia, such as Salihović et al. (Salihović et al., 2018), discovered dementia predictors and deficits in multiple cognitive functions in vascular cognitive diseases. A recent study by Nyholm et al. (Nyholm et al., 2023) used machine learning to identify the risk factors for early prediction of dementia based on sleep disturbances in older adults. Wang et al. (Wang et al., 2019), examined the association between the difference in expected and chronological brain age and the development of dementia in a large population-based cohort of adult and older individuals using a deep learning model. According to the outcomes, the difference between expected and historical brain ages is a biomarker associated with dementia risk. It could be used as an additional biomarker for dementia threat assessment. Shigemizu (Shigemizu et al., 2019) constructed an optimal risk prediction model based on several ML methods, including penalized regression, RF, support vector machines, and gradient boosting decision tree, employing blood miRNA expression data from 478 Japanese adults. Ryu et al. (Ryu et al., 2020), address the subject of population aging and the growth of geriatric illnesses, notably dementia, which is lethal to the daily activities of the elderly. The authors provide a dementia prediction paradigm based on XGBoost, an ML algorithm that uses the derived variable extraction method. By extracting variable significance from traditional independent variables, they use gradient boosting to generate derived variables. The obtained variables' findings are utilized to perform variable significance analysis, leading to the development of a Top-N group. Hyper-parameter alteration is used to achieve the best efficiency compatible with the data features for each Top-N group. The authors compare the performance of the proposed model to that of current ML classification methods. The effects of biomarker-based dementia risk estimation on quality of life (QoL) in mild cognitive impairment (MCI) patients and their immediate relatives have not been adequately examined.

Furthermore, Rostamzadeh (Rostamzadeh et al., 2021) provided empirical information on the effects of prediction on QoL and developed an ethical and legal framework for biomarker-based dementia risk assessment in MCI. Kühnel et al. (Kühnel et al., 2021), conducted a study to develop and validate a continuous biomarker-based model for estimating an individual's cognitive level at any point in the future. Ghazal et al. (Ghazal et al., 2022), used ML classifiers to predict cancer, dementia, and diabetes using different datasets. Their proposed approach for multiclass classification used support vector machines (SVM) and K-nearest neighbor (KNN) ML algorithms to forecast three circumstances and compare the accuracy of these tactics.

For reliable dementia prediction, Javeed et al. (Javeed et al., 2022b), presented a hybrid diagnostic system based on ML algorithms. In their proposed method, they constructed an autoencoder that extracted features from the dataset, and an ensemble learning model was used for classification. In another study, Javeed et al. (Javeed et al., 2023d), presented a pair of automated diagnostic systems that use genetic algorithms for feature selection. At the same time, artificial neural networks (ANN) and deep neural networks (DNN) are used for dementia classification. Based on a genetic algorithm and a deep neural network, the suggested model had the highest accuracy of 93.36%, sensitivity of 93.15%, and specificity of 91.59%. Moreover, ML models tend to favor the majority class in the dataset. To solve this problem, Javeed et al. (Javeed et al., 2023e) proposed a diagnostic system for the early detection of dementia using an adaptive synthetic sampling technique (ADASYN) to solve the problem of imbalance in the dataset. They proposed novel feature extraction techniques, namely, feature extraction batteries (FEB) and optimized support vector machines (SVM) using radical basis functions (RBF), for dementia classification. The grid search method was used to calibrate the SVM hyperparameters. Their proposed model (FEB-SVM) increased the dementia prediction accuracy of the standard SVM by 6%. The proposed model (FEB-SVM) achieved a training accuracy of 98.28% and a test accuracy of 93.92%. The proposed approach achieved a precision of 91.80%, a recall of 86.59%, and an F1 score of 89.12%. M. A. Maito et al. (Maito et al., 2023) presented a fully automated computational approach based on classical statistical and machine learning methods for dementia prediction by identification of risk factors for dementia. The classification of Alzheimer's disease (AD) and frontotemporal dementia (FTD) patients was shown to be accurate based on the results. With an accuracy of 0.91%, a machine learning model generated the optimal values to distinguish AD patients from FTD patients. M. Bucholc et al. (Bucholc et al., 2023) proposed a novel prognostic machine learning (ML) framework to identify mild cognitive impairment (MCI) patients who are susceptible to dementia by utilizing longitudinal data encoded in effective, affordable, and non-invasive markers. By their proposed method, RF and ensemble models had the highest reported accuracy, at 87.5% and 86.8%, respectively. Another study by Javeed et al. (Javeed et al., 2023c) aims to thoroughly evaluate the automated diagnostic systems previously presented by the researchers based on ML, using multiple data modalities such as images, medical variables, and audio data.

2 Materials and methods

2.1 Dataset description

The data for this study was obtained from the Swedish National Study on Ageing and Care (SNAC). In 1999, the Swedish Ministry of Social Affairs launched and funded a nationwide initiative to monitor and analyze the Swedish elderly care system. Four longitudinal, individual-based data-gathering projects characterizing the aging process and embracing the whole care system have been launched to accomplish these goals. The Swedish National Study on Aging and Care (SNAC) was the name given to this initiative. The SNAC is a long-running organization that collects multimodal data from Sweden's aging population to offer reliable, efficient, and long-term data sets for aging research (Lagergren et al., 2004).

The SNAC program was developed to assess the quality of healthcare provided to older adults in multiple ways. Medical records, social variables, lifestyle factors, metacognitive data, and physical examination are only a few topics covered by SNAC's databases. As a result, variables were chosen from the SNAC databases (Blekinge, Kungsholmen, and Skåne) based on previously published research in eight areas, including demographics, social factors, lifestyle, medical history, physical exam, biochemical testing, psychological exam, and evaluation of various health devices (Arvanitakis et al., 2019; Yu et al., 2020). In total, 75 variables were selected from the areas mentioned above. Table 1 provides the overview of selected variables from the SNAC databases.

For this investigation, we acquired data from three SNAC facilities (Blekinge, Kungsholmen, and Skåne). Forty-three thousand forty data samples were collected, with 3,461 coming from SNAC-Kungsholmen, 7,304 from SNAC-Blekinge, and 32,275 from SNAC-Skåne. The primary purpose of the data collection was to integrate and harmonize the SNAC data from the three sites. The dataset was standardized by following the harmonization rule for the dataset's variables. In the acquired dataset, there are 18,312 men and 24,728 women. Only 819 of the 18,312 males and 1,059 of the 24,728 females are affected by dementia. Table 2 displays the demographic information for the participant group. Because this study is based on broad criteria, no boundaries exist between urban and rural areas. Subjects excluded from this study based on the following criteria: Participants with dementia at baseline; participants with missing data for the outcome variable (dementia diagnosis); participants with more than 10% missing data for the entry variable; participants who died before the 10-year study or were diagnosed with advanced dementia were excluded (Figure 1) group.

2.2 Methodology

Real-world datasets come in a variety of sizes and forms. As a result, their nature imposes several significant limits on both learning models and feature selection techniques (Liu and Yu, 2005). Sample sizes and feature counts for datasets may be substantial, and problems with redundant, noisy, multivariate, and nonlinear scenarios may also arise. As a result, most currently used methods need help to solve these issues. Additionally, there is no such thing as "the best feature selection method" in general, which makes it challenging for users to choose one way over another. A user is expected to comprehend the

TABLE 1 Overview of selected variables (features).

Variable_Category	Variable_Names	Sum
Demographic	Gender, Age	02
Social	Support Network, Education, Loneliness, Religious Belief, Social Network, Voluntary Association, Religious Activities	07
Lifestyle	Physical-Workload, Past Smoker, Physically Demanding Activities, Present Smoker, Alcohol Quantity, Leisure Activities, Number of Cigarettes a Day, Social Activities, Work Status, Alcohol Consumption, Light Exercise	11
Medical History	High Blood Pressure, Myocardial Infarction, Head Trauma, Arrhythmia, Sleep Apnea, TIA/RIND, Diabetes Type 1, Stroke, Thyroid Disease, Family History of Importance, Cancer, Epilepsy, Number of Medications, Diabetes Type 2, Parkinson's Disease, Other Psychiatric Diseases, Snoring, Atrial Fibrillation, Hip Fracture, Cardiovascular Ischemia, Heart Failure, Developmental Disabilities, Depression	22
Physical Examination	Blood Pressure on the Right Arm, Number of Teeth, Body Mass Index (BMI), Assessment of Rising from a Chair, Heart Rate Sitting, Heart Rate Lying, Hand Strength in Right Arm in a 10s Interval, Hand Strength in Left Arm in a 10s Interval, Single-Leg Standing with Right Leg, Single Leg Standing with Left Leg, Pain in the last 4 weeks, Dental Prosthesis, Feeling of Safety from Rising from a Chair	13
Biochemical Test	C-Reactive Protein Analysis, Hemoglobin Analysis	02
Psychological	Sense of Identity, Memory Loss, Personality Change, Memory Decline, Abstract Thinking, Memory Decline 2	06
Health Instruments	Backwards Digit Span Test, Digit Span Test, Livingston Index, Sense of Coherence EQ5D Test, Instrumental Activities of Daily Living, Mini-Mental State Examination, Activities of Daily Living, Comprehensive Psychopathological Rating Scale, Physical Composite Score of the SF-12 Health Survey, Clock Drawing Test, Mental Composite Score of the SF-12 Health Survey	12

TABLE 2 Summary of samples population.

Age_Group	Female		Male	
	+ve	-ve	+ve	-ve
60-70	334	2,471	319	2,414
70-80	203	6,372	192	5,603
80-90	88	5,701	82	4,411
90+	434	9,125	226	5,065
Total	1,059	23,669	819	17,493

technical specifics of the various algorithms and have a thorough understanding of each dataset's domain and features to make the best decision (Tuv et al., 2009).

Many businesses today depend on machine learning techniques to extract meaningful data and expertise from escalating large datasets. For classification problems, feature selection methods are employed to mine the most relevant features in a feature space (Ali et al., 2019b; Akbar et al., 2020; Javeed et al., 2020; 2023d). In this sense, we employed a feature ranking method based on the statistical technique F-score. The F-score based on the feature ranking model (Chen and Lin, 2006) measures the discriminating between two sets of real numbers. If the number of samples associated with healthy participants is v_+ and the number of samples of the patient group is v_- for a specific dataset with l_i , $i = 1, 2, 3, \dots, n$ occurrences, the F-score of the n th feature is determined as:

$$M_1 = \frac{1}{(v_+ - 1)} \sum_{i=1}^{v_+} \left(\bar{l}_{i,n}^{(+)} - \bar{l}_n^{(+)} \right)^2 \tag{1}$$

$$M_2 = \frac{1}{(v_- - 1)} \sum_{i=1}^{v_-} \left(\bar{l}_{i,n}^{(-)} - \bar{l}_n^{(-)} \right)^2 \tag{2}$$

From Eqs 1 and 2, we get Eq. (3)

$$H(n) = \frac{\left(\bar{l}_n^{(+)} - \bar{l}_n \right)^2 + \left(\bar{l}_n^{(-)} - \bar{l}_n \right)^2}{M_1 + M_2} \tag{3}$$

Here, \bar{l}_n , \bar{l}_n^+ and \bar{l}_n^- are means of the n^{th} feature of the positive and negative samples in the complete dataset. Furthermore, \bar{l}_n is n^{th} feature of the i th negative sample, and \bar{l}_n^+ is i^{th} the positive sample of the n^{th} feature. A further distinction between the positive and negative sets is made in 3's numerator, while the denominator designates the singular value within each of the two sets (Akay, 2009). The F-score value is inversely correlated with a feature's discriminative capacity. Following the statistical model-based ranking of features by their F-score, we must choose a threshold for the F-score, meaning only those features will be chosen if their F-score exceeds the threshold. In the current work, we use the hybrid grid search algorithm (HGSA) to find the ideal threshold to yield an ideal subset of the selected features. The extracted feature subset is supplied to the proposed ensemble model for classification.

This study presents a framework using the F-score statistical method for feature selection and a voting ensemble classifier based on five machine learning (ML)-based models such as DT, SVM, NB, LR, and RF (Naseem et al., 2022) as shown in Figure 2. The complicated voting process determines the class with the most votes as the expected outcome (Bauer and Kohavi, 1999; Zhang et al., 2014). The proposed voting classifier used hard voting scheme for expected outcome. The next part also briefly explains each algorithm employed in our investigation.

2.3 Decision tree (DT)

DT comprises massive classifications of samples into specific categories Salzberg (1994). The utilization of specimens is made possible by patterns that combine nominal and numerical information to provide precise group descriptions. These indicators are then represented as models, producing decision

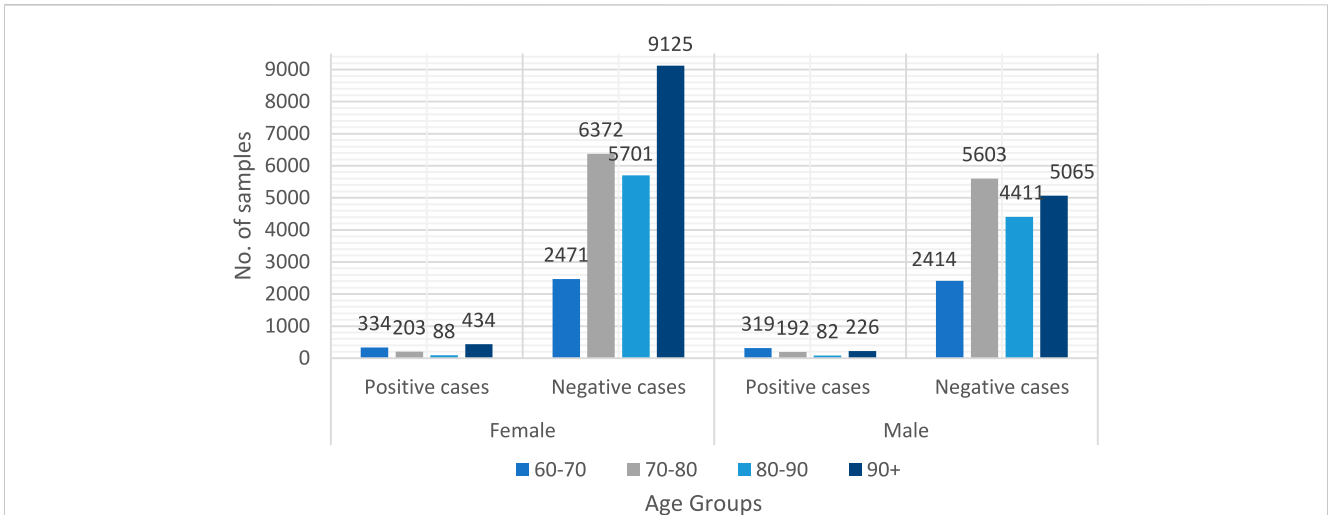


FIGURE 1
Samples overview in the collected dataset.

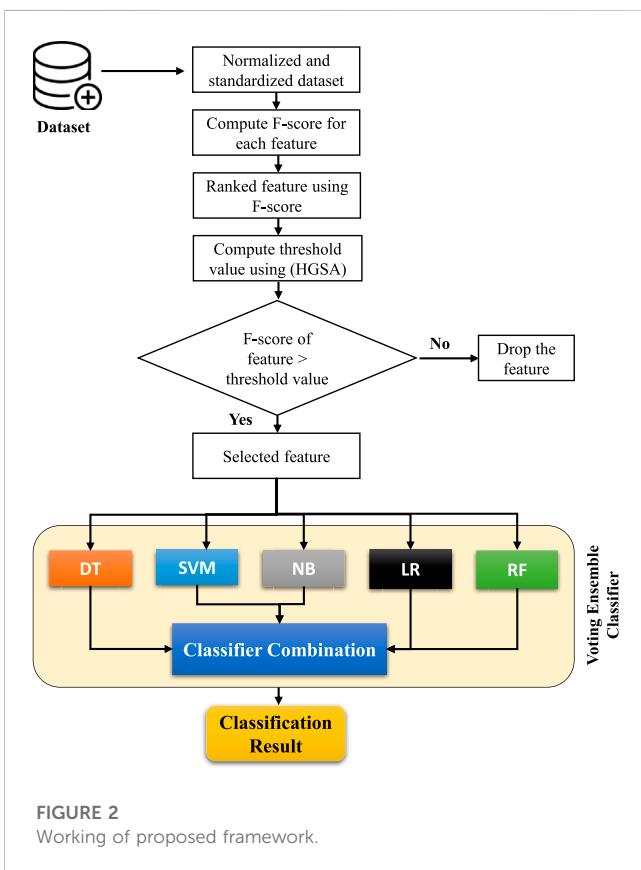


FIGURE 2
Working of proposed framework.

frameworks or collections of if-then procedures that may be used to discriminate new samples, emphasizing the importance of creating clear and precise designs. The C4.5 calculus chooses the test that extracts the most data from a group of specimens without confining themselves to evaluating a single characteristic. They then apply equations based on theoretical data to estimate the ‘goodness’ of the test. Dealing with the overfitting and unknown values problem is

DT’s main drawback. Unknown values are a problem that can be solved using the DT C4.5 approach, especially since samples with unknown values are usually ignored. A classifier that classifies every sample in the training set might not be as efficient as a DT. C4.5 implements an error rate-based pruning process for all subtrees to get around this. This method removes the subtree when the computed error is high. This method is more efficient and yields superior outcomes (Moslehi et al., 2022).

2.4 Support vector machines (SVM)

SVM is a supervised machine learning algorithm that constructs a linear discriminant function by using support vectors, which are an adequate number of samples. SVM resolved the linear constraints (Awad et al., 2015). A maximum hyperplane margin can be seen by partitioning the SVM data linearly into two classes. After choosing the suitable mapping, the new samples are linearly fitted or seem linearly separable in the high-level plane. The classification is done by graphing the best hyperplane, which can categorize the data objects according to their features (Javeed et al., 2023e). The following is a representation of the hyperplane:

$$\omega \cdot \rho + \beta = 0 \tag{4}$$

where the offset is represented by β and ω by the plane’s normal vector. Following the formation of the hyperplane, classifications based on the input vectors can be made. The following is a representation of the prediction hypothesis:

$$H(x_i) = \begin{cases} +1 & \text{if } \omega \cdot x + \beta \geq 0 \\ -1 & \text{if } \omega \cdot x + \beta < 0 \end{cases} \tag{5}$$

When the input point is either above or on the hyperplane, it is categorized as positive, or $\omega \cdot x + \beta \geq 0$. If it is under the hyperplane, it is defined as negative, or $\omega \cdot x + \beta < 0$. In the SVM, improvements are

typically made by increasing the hyperplane's separation from the support vectors.

2.5 Naive Bayes (NB)

The NB method is proposed using the Bayes theorem (Yager, 2006). The Bayes theorem and the precise processes can be used to revise the NB classifier in the following ways (Zheng et al., 2018). We conclude that a training set of examples S exists. These specimens bear group markings. The names of the groupings are G_1, G_2, \dots, G_n . Every specimen is an agent with n dimensions, denoted by formula $D = d_1, d_2, \dots, d_n$. It claims that because D has n dimensions, it has n characteristics. If the likelihood that group i depends on a given specimen, D , is higher than the likelihood that each of the other groups depends on D , then D is projected to belong to group G_i , as given:

$$P(G_i|D) > P(G_j|D) \text{ for } 1 \leq j \leq k_j = i \quad (6)$$

$P(G_i|D)$ is determined by the Bayes' Theorem as follows:

$$P(G_i|D) = P(D|G_i)P(G_i/P(D)) \quad (7)$$

2.6 Logistic regression (LR)

LR is a supervised classification technique that predicts a category based on input attributes. LR is a predictive approach that makes predictions using probability values ranging from 0 to 1. As a result, an S-curve, also known as the sigmoid function, is formed. If the anticipated probability value exceeds a certain threshold, it is classified as positive; otherwise, it is classified as negative (Boateng and Abaye, 2019). The LR can be calculated using the formula below. A straight-line's equation for LR is:

$$N = \alpha_0 + \alpha_1\beta_1 + \alpha_2\beta_2 + \dots + \alpha_n\beta_n \quad (8)$$

If the value of N is between 0 and 1, divide N by $1 - N$.

$$\left[\frac{N}{1 - N} \right] = \alpha_0 + \alpha_1\beta_1 + \alpha_2\beta_2 + \dots + \alpha_n \quad (9)$$

2.7 Random forest (RF)

From the feature vectors, the RF approach generates n -tree bootstrap samples. Each sample is used to build a classification tree that has not been trimmed. Each tree node evaluates a random collection of 'F' features and chooses the optimal split from them (Javeed et al., 2019; 2022a). This algorithm predicts the class of new, unknown data by aggregating the predictors of n trees using the majority voting technique (Paul et al., 2018).

Two hyperparameters are crucial for the classification job by RF model, such as D , the depth of each tree, and E , the number of trees making up the forest (Liu et al., 2015; Javeed et al., 2023f). In order to guarantee the enhanced performance of the random forest model, the best E and D were found in this study using the random search

algorithm (RSA). In addition, a new sample is added, and an RF model is created. In the same manner, the decision tree determines and evaluates the new sample type. The final classification of a sample can be ascertained using the total number of votes cast in the decision tree within the forest. The bootstrap technique builds the RF formation trees from repeated samples by using training data. To apply the replacement method for model interactions, bootstrap is a straightforward and practical solution (Aprilliani and Rustam, 2018). Using bootstrap random sampling, a predetermined number of samples are taken from the training set. The number of samples that were extracted, the number of samples that were returned to the training set, and the number of bootstrap samples that were produced. It is also possible that the extracted samples will be re-sampled once the training set is returned. Thus, it is better to sample the previously extracted samples after storing them.

3 Experimental results

3.1 Evaluation metric

Several validation techniques, including holdout validation and cross-validation, are used in data mining and machine learning to assess how well a developed ML model performs. The cross-validation method has certain advantages over the holdout method, such as each partition of the data set used for training and testing of the ML models (Javeed et al., 2023b). Hence, to validate the performance of the proposed method for the prediction of dementia, we employed cross-validation schemes (Ali et al., 2019b; Liu et al., 2023b; Saleem et al., 2023).

The performance of the newly proposed method is assessed on several evaluation metrics, such as accuracy, sensitivity, specificity, and area under the curve (AUC), by employing the receiver operating characteristic curve (ROC) (Ali et al., 2019a; Liu et al., 2023a). Accuracy is given as follows:

$$Accuracy = \frac{\text{No. of correct prediction}}{\text{Total samples}} \quad (10)$$

Sensitivity and specificity are defined as follows:

$$Sensitivity = \frac{TP}{TP + FN} \quad (11)$$

$$Specificity = \frac{TN}{TN + FP} \quad (12)$$

When evaluating a classification model, TN represents true negatives, FN represents false negatives, FP represents false positives, and TP represents true positives.

The performance of predictive models must be measured using statistical analysis. In our statistical analysis of binary classification, we employed the Matthews correlation coefficient (MCC). The test's accuracy is ascertained using MCC, a variable with values ranging from -1 to 1 . Where -1 represents poorer predictions and 1 indicates exact predictions. The mathematical formulation of MCC is given as follows:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (13)$$

TABLE 3 Performance of baseline ML Classifier.

ML models	Accuracy	Sensitivity	Specificity	MCC
LR	88.90	90.76	87.85	0.6699
NB	87.20	95.68	22.74	0.5423
LDA	82.90	88.14	72.49	0.6260
AdaBoost	89.50	92.51	84.52	0.6844
kNN	89.00	91.42	88.00	0.6733
DT	87.10	86.41	90.57	0.6577
SVM	85.20	88.00	82.36	0.6468
RF	88.70	92.87	84.14	0.6600

TABLE 4 Performance of proposed model.

Sub_Fe.	Accuracy	Sensitivity	Specificity	MCC
02	94.64	95.64	48.50	0.6513
04	92.10	96.42	79.00	0.5944
06	95.62	96.82	92.50	0.6852
07	95.05	96.91	85.45	0.6600
08	98.12	98.00	93.00	0.7515
09	98.25	97.44	95.74	0.7535
10	97.90	96.77	94.40	0.7412
12	98.00	97.69	95.00	0.7500

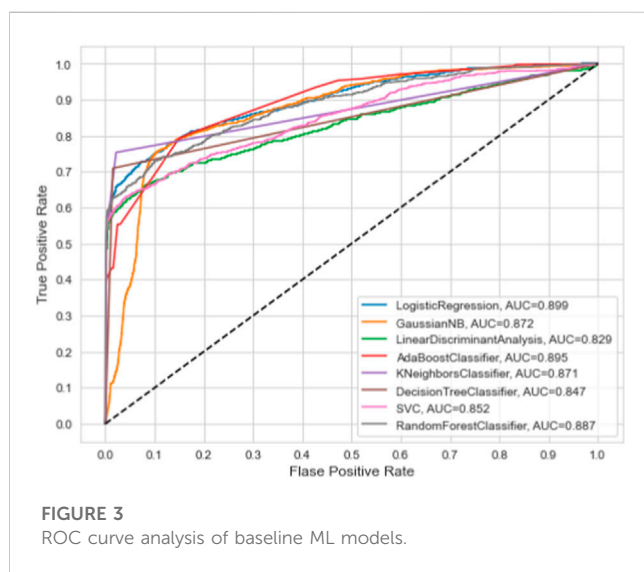


FIGURE 3
ROC curve analysis of baseline ML models.

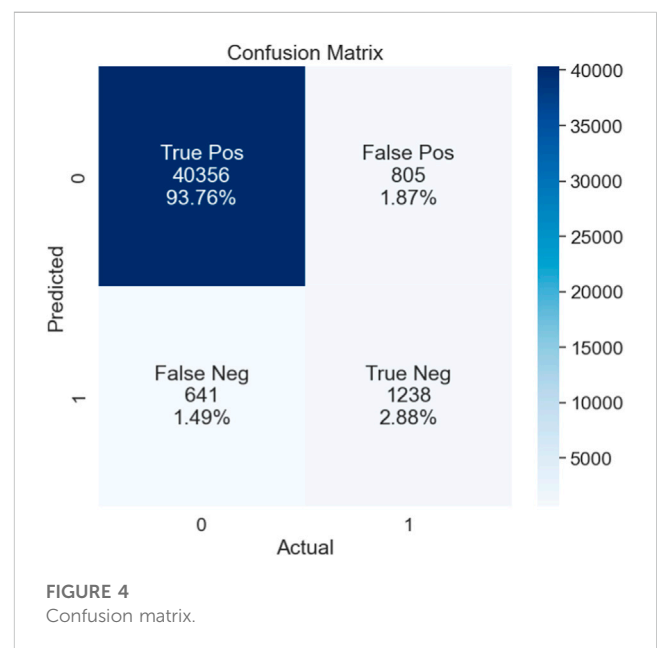


FIGURE 4
Confusion matrix.

4 Results

4.1 Experiments no: 1 performance of baseline ML models

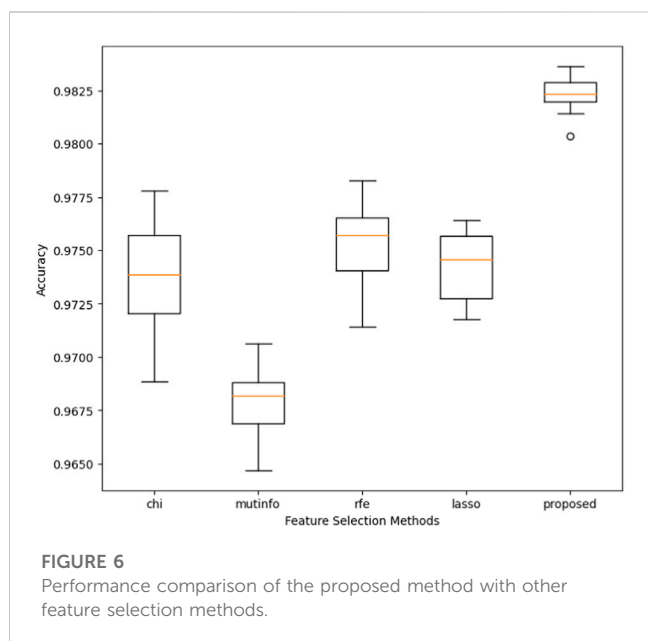
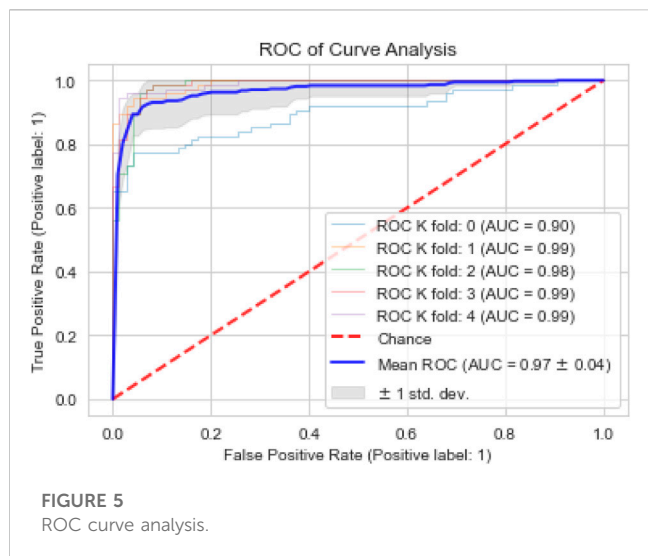
This section summarizes the findings and presents the baseline models against which our suggested model was compared. We compare the proposed model’s performance to numerous machine learning classifiers and other approaches. All of the algorithms in this study were performed with their default parameters. Table 3 presents the results of baseline ML models on the given dataset. Furthermore, we have also employed the ROC curve to validate the performance results of based line ML models, as given in Figure 3. The ML model’s performance was evaluated by utilizing all the features present in the dataset. Table 3 and Figure 3 show that the highest accuracy and AUC are achieved through the LR model by using all the dataset features. While the worst performance for the prediction of dementia in terms of accuracy 82.90% is given by LDA.

In the next phase, we validate the performance of the proposed method, where useful features are selected from the dataset by employing the statistical method F-score. We constructed a voting classifier based on DT, SVM, NB, LR, and RF for the classification job.

We have extracted different sizes of subsets of features from the dataset and measured their performance using a voting classifier. Table 4 presents the performance of the proposed model on different subsets of features. The performance of the proposed model is measured in terms of accuracy, sensitivity, specificity, and MCC.

4.2 Experiments no: 2 performance of proposed model

In this section, we evaluate the performance of the proposed model, where features from the dataset are selected based on the statistical method (F-score) and the ensemble voting classifier performs the classification task. We have stacked five different ML models for the voting classifier, i.e., DT, SVM, NB, LR, and RF. To validate the efficiency of the newly developed method for the prediction of dementia, we employed a cross-validation scheme ($k = 5$) to avoid the problem of model overfitting. Table 4 presents the performance of the proposed model based on accuracy, sensitivity, specificity, and MCC by using a subset of features (Sub_Fe) extracted



by the F-score. Table 4 shows that the highest accuracy achieved by the proposed model was 98.25%, using only nine features from the dataset.

The performance of the proposed model is also assessed based on a confusion matrix, as seen in Figure 4. In binary classification problems, the efficiency of ML models is often validated using the ROC curve. A larger area under the curve (AUC) indicates a more efficient model. We also used the ROC curve to evaluate our proposed model's performance. Figure 5 shows that our model achieved the highest AUC of 97% through the cross-validation ($k = 5$) scheme.

4.3 Experiments no: 3 performance of other feature selection methods

In this section, we conducted the experiment where the performance of different feature selection methods, i.e., chi-

square test, mutual information (mutinfo), recursive feature elimination (RFE), least absolute shrinkage, and selection operator (lasso), was evaluated with a constructed voting classifier for the classification.

Furthermore, we compared the performance of the proposed model with the feature mentioned above selection algorithms. Figure 6 presents the performance comparison regarding the proposed model's accuracy and the four state-of-the-art feature selection techniques (chi-square, mutinfo, RFE, and lasso). The accuracy is measured through the cross-validation scheme, where the value of k was set to 10. The proposed model achieved the highest accuracy of 98.25%, while mutinfo, along with the constructed voting classifier, obtained the lowest accuracy of 96.75%.

5 Discussion

For this work, we developed a large dataset from the Swedish National Study on Ageing and Care (SNAC) for the early prediction of dementia and its risk factors. SNAC is a cohort-based study collecting data from older Swedish adults since 2002. For this study, we gathered data from three distinct locations in Sweden (Blekinge, Kungsholmen, and Skåne). In total, 43,040 data samples were collected, comprising 75 features for each sample. The description of selected features for this study is given in Table 2, where features belong to eight different categories such as lifestyle (11), demography (02), social (07), medical history (22), health instruments (12), biochemical tests (02), psychological (06) and physical examination (13). We employed data harmonization rules to integrate the data from three SNAC centers. After data collection, we clean the dataset by performing data standardization and normalization techniques. This study aimed to design a diagnostic system that can predict the early onset of dementia in older adults and detect the risk factors that cause dementia. For this purpose, we proposed a hybrid diagnostic system based on statistical methods and machine learning techniques. From feature space, highly significant features are selected through statistical method (F-score). We designed an ensemble voting classifier based on 5 ML classifiers (DT, NB, SVM, LR, and RF) for the classification task. The proposed system generated a subset of significant features, which were tested by constructed voting classifiers to accurately predict dementia. As dementia is rare, the number of instances of dementia in comparison to the healthy instances is less. The ML models tend to overfit due to the majority class in the dataset.

To avoid this problem, we deployed several evaluation metrics based on a cross-validation scheme to assess the efficiency of the newly constructed model. From Table 4, it is evident that the proposed model achieved the highest accuracy of 98.25% by using only nine features from the dataset. Table 5 describes nine highly significant features selected by the proposed model that help to predict dementia. Most of the features selected by the proposed model for the prediction of dementia are related to subject health conditions such as heart disease, TIA/RIND, head trauma, psychosis, and arrhythmia. While few are associated with social interaction, such as receiving low assistance regarding personal

TABLE 5 Significant features selected by the proposed model.

Feature.	Description
E18	Heart disease
E19	TIA/RIND*
E23	Head Trauma with loss of consciousness/syncope
E44	Psychosis
E58_C	ECG - Arrhythmia
F155	I happily, and often, entertain myself with abstract ideas or theories
FP10	Mental Rotations Test- Total number of correct responses (0–10)
H27	Receive not enough assistance regarding personal care
BMI	Body Mass Index

*TIA/RIND: transient ischaemic attack reversible ischaemic neurological deficit.

care and entertaining with abstract ideas, One selected feature (the mental rotation test) is related to psychology, where subjects with low scores are more prone to dementia. Subjects with a massive BMI also have dementia in older age.

The constructed voting classifier also evaluates the performance of other feature selection (chi square, ref, mutinfo, lasso) methods. Figure 6 shows that the proposed model obtained the highest accuracy compared to the rest of the feature selection methods.

Furthermore, we also evaluated the performance of the conventional ML models using all features from the dataset. Table 3 shows that the AdaBoost obtained the highest accuracy of 89.59% while LDA achieved the lowest accuracy of 82.90%.

Furthermore, we have also compared the performance of the proposed method’s classification accuracy to other methods in the literature using the dementia dataset. Table 6 provides a succinct summary of these approaches. The newly proposed method outperforms the eighteen recently proposed methods in terms of accurate prediction of dementia. Furthermore, the presented framework has two components, and they work in sequential coordination. The first component is used for the feature selection from the dataset, and the second component is employed for the classification. Thus, the computational complexity of the proposed framework is $\mathcal{O}(\log n)$.

Although the newly proposed method has shown evident performance in terms of accuracy, there are a few concerns that need to be addressed in future research work. One of the limitations of this study is that it uses only electronic health record data. Therefore, in the future, researchers should focus on multimodality datasets for the prediction of dementia. Hybrid diagnostic systems are complex in nature, especially in terms of computational and time complexity; thus, novel diagnostics should be developed in the future based on machine learning and deep learning that are simple and easily integrated into the real-world scenario.

TABLE 6 The proposed method’s classification accuracy compared to other methods in the literature using the EHR of dementia.

S.No	Study	Method	Validation	Accuracy (%)
01	Cho and Chen (2012)	PNNs	Holdout	83.00
02	Gurevich et al. (2017)	SVM	K-Fold	89.00
03	Stamate et al. (2018)	Gradient Boosting	K-Fold	88.00
04	Visser et al. (2019)	XGBoost + RF	K-Fold	88.00
05	Dallora et al. (2020)	DT	K-Fold	74.50
06	Karaglani et al. (2020)	RF	Holdout	84.60
07	Ryzhikova et al. (2021)	ANN + SVM	K-Fold	84.00
08	Salem et al. (2021)	RF	K-Fold	88.00
09	Garcia-Gutierrez et al. (2022)	GA	K-Fold	84.00
10	Hsiu et al. (2022)	MLP	K-Fold	70.32
11	Shahzad et al. (2022)	SVM	K-Fold	71.67
12	Javeed et al. (2022b)	Autoencoder + Adaboost	Holdout	90.23
13	Bucholc et al. (2023)	Ensemble Model	K-Fold	87.50
14	Maito et al. (2023)	Anova test + SVM	K-Fold	91.00
15	Nyholm et al. (2023)	Gradient Boosting	K-Fold	92.90
16	Javeed et al. (2023d)	GA + DNN	Holdout	93.36
17	Javeed et al. (2023e)	FEB + SVM	Holdout	93.92
18	Javeed et al. (2023a)	LDA + DT	Holdout	97.77
19	Proposed Method (2023)	F-Score + Voting Classifier	K-Fold	98.25

6 Conclusion

This study presented a hybrid diagnostic system that helps predict dementia in its early stages by using the medical health records of older adults. The newly developed diagnostic system consists of two modules: statistical models and ensemble ML models. There are 75 features in the dataset. To eliminate the irrelevant features from the feature space, we deployed a statistical (F-score) model that helps construct a subset of useful features from the dataset. We developed a voting classifier based on DT, SVM, NB, LR, and RF for the classification. The extracted subset of features from the first module of the developed diagnostic system is fed into the second module for the classification of dementia. To assess the performance of the developed diagnostic system, we employed a cross-validation scheme to overcome the problem of ML model overfitting. Various evaluation metrics were adopted to rigorously validate the efficiency of the developed diagnostic system, such as accuracy, sensitivity, specificity, ROC, and MCC. The experimental results show that the proposed diagnostic system achieved an accuracy of 98.25%, with a sensitivity of 97.44%, a specificity of 95.74%, and an MCC of 0.7535. Furthermore, the performance of the proposed diagnostic system is also compared with the baseline ML models as well as other state-of-the-art feature selection methods. In this context, four different feature selection methods, such as chi square, mutinfo, ref, and lasso, were tested along with the constructed voting ensemble classifier for the classification. Regarding accuracy, the proposed model outperformed the rest of the feature selection methods and baseline ML models.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

References

- Ahiskali, M., Polikar, R., Kounios, J., Green, D., and Clark, C. M. (2009). "Combining multichannel erp data for early diagnosis of alzheimer's disease," in 2009 4th International IEEE/EMBS Conference on Neural Engineering, Antalya, Turkey, April, 2009, 522–525.
- Akay, M. F. (2009). Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Syst. Appl.* 36, 3240–3247. doi:10.1016/j.eswa.2008.01.009
- Akbar, W., Wu, W.-p., Saleem, S., Farhan, M., Saleem, M. A., Javeed, A., et al. (2020). Development of hepatitis disease detection system by exploiting sparsity in linear support vector machine to improve strength of adaboost ensemble model. *Mob. Inf. Syst.* 2020, 1–9. doi:10.1155/2020/8870240
- Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A., and Khan, J. A. (2019a). An automated diagnostic system for heart disease prediction based on χ^2 statistical model and optimally configured deep neural network. *Ieee Access* 7, 34938–34945. doi:10.1109/access.2019.2904800
- Ali, L., Zhu, C., Golilarz, N. A., Javeed, A., Zhou, M., and Liu, Y. (2019b). Reliable Parkinson's disease detection by analyzing handwritten drawings: construction of an unbiased cascaded learning system based on feature selection and adaptive boosting model. *Ieee Access* 7, 116480–116489. doi:10.1109/access.2019.2932037
- Aprilliani, U., and Rustam, Z. (2018). "Osteoarthritis disease prediction based on random forest," in 2018 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Yogyakarta, Indonesia, October, 2018, 237–240.
- Arvanitakis, Z., Shah, R. C., and Bennett, D. A. (2019). Diagnosis and management of dementia: review. *Jama* 322, 1589–1599. doi:10.1001/jama.2019.4782
- Awad, M., Khanna, R., Awad, M., and Khanna, R. (2015). Support vector machines for classification. *Effic. Learn. Mach. Theor. Concepts, Appl. Eng. Syst. Des.*, 39–66. doi:10.1007/978-1-4302-5990-9_3
- Basheer, S., Bhatia, S., and Sakri, S. B. (2021). Computational modeling of dementia prediction using deep neural network: analysis on oasis dataset. *IEEE access* 9, 42449–42462. doi:10.1109/access.2021.3066213
- Bauer, E., and Kohavi, R. (1999). An empirical comparison of voting classification algorithms: bagging, boosting, and variants. *Mach. Learn.* 36, 105–139. doi:10.1023/a:1007515423169
- Boateng, E. Y., and Abaye, D. A. (2019). A review of the logistic regression model with emphasis on medical research. *J. data analysis Inf. Process.* 7, 190–207. doi:10.4236/jdaip.2019.74012
- Bucholc, M., Titarenko, S., Ding, X., Canavan, C., and Chen, T. (2023). A hybrid machine learning approach for prediction of conversion from mild cognitive impairment to dementia. *Expert Syst. Appl.* 217, 119541. doi:10.1016/j.eswa.2023.119541
- Chen, Y.-W., and Lin, C.-J. (2006). Combining svms with various feature selection strategies. *Feature Extr. Found. Appl.*, 315–324. doi:10.1007/978-3-540-35488-8_13
- Cho, P.-C., and Chen, W.-H. (2012). "A double layer dementia diagnosis system using machine learning techniques," in International Conference on Engineering Applications of Neural Networks, London, UK, September, 2012, 402–412.
- Creavin, S. T., Wisniewski, S., Noel-Storr, A. H., Trevelyan, C. M., Hampton, T., Rayment, D., et al. (2016). Mini-mental state examination (mmse) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Cochrane Database Syst. Rev.* 2016, CD011145. doi:10.1002/14651858.cd011145.pub2
- Dallora, A. L., Minku, L., Mendes, E., Rennemark, M., Anderberg, P., and Sanmartin Berglund, J. (2020). Multifactorial 10-year prior diagnosis prediction model of dementia. *Int. J. Environ. Res. public health* 17, 6674. doi:10.3390/ijerph17186674

Author contributions

AJ: Methodology, Writing—original draft. PA: Conceptualization, Writing—review and editing. AG: Writing—review and editing. AN: Writing—original draft. SE: Investigation, Visualization, Writing—review and editing. JB: Supervision, Writing—original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. Open access funding provided by Blekinge Institute of Technology. The first author's learning process was supported by the National E-Infrastructure for Aging Research (NEAR), Sweden. NEAR is working on improving the health condition of older adults in Sweden.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer NA declared a shared affiliation with the author AB to the handling editor at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Duchesne, S., Caroli, A., Geroldi, C., Barillot, C., Frisoni, G. B., and Collins, D. L. (2008). Mri-based automated computer classification of probable ad versus normal controls. *IEEE Trans. Med. Imaging* 27, 509–520. doi:10.1109/tmi.2007.908685
- García-Gutiérrez, F., Delgado-Alvarez, A., Delgado-Alonso, C., Díaz-Alvarez, J., Pytel, V., Valles-Salgado, M., et al. (2022). Diagnosis of alzheimer's disease and behavioural variant frontotemporal dementia with machine learning-aided neuropsychological assessment using feature engineering and genetic algorithms. *Int. J. Geriatric Psychiatry* 37. doi:10.1002/gps.5667
- Ghazal, T. M., Al Hamadi, H., Umar Nasir, M., Gollapalli, M., Zubair, M., Adnan Khan, M., et al. (2022). Supervised machine learning empowered multifactorial genetic inheritance disorder prediction. *Comput. Intell. Neurosci.* 2022, 1–10. doi:10.1155/2022/1051388
- Gurevich, P., Stuke, H., Kastrop, A., Stuke, H., and Hildebrandt, H. (2017). Neuropsychological testing and machine learning distinguish alzheimer's disease from other causes for cognitive impairment. *Front. Aging Neurosci.* 9, 114. doi:10.3389/fnagi.2017.00114
- Hsiu, H., Lin, S.-K., Weng, W.-L., Hung, C.-M., Chang, C.-K., Lee, C.-C., et al. (2022). Discrimination of the cognitive function of community subjects using the arterial pulse spectrum and machine-learning analysis. *Sensors* 22, 806. doi:10.3390/s22030806
- Iadecola, C. (2016). Vascular and metabolic factors in alzheimer's disease and related dementias: introduction. *Cell. Mol. Neurobiol.* 36, 151–154. doi:10.1007/s10571-015-0319-y
- Javeed, A., Ali, L., Mohammed Seid, A., Ali, A., Khan, D., and Imrana, Y. (2022a). A clinical decision support system (cdss) for unbiased prediction of caesarean section based on features extraction and optimized classification. *Comput. Intell. Neurosci.* 2022, 1–13. doi:10.1155/2022/1901735
- Javeed, A., Berglund, J. S., and Anderberg, P. (2023a). Empowering dementia diagnosis: a machine Learning-Driven automated system. *Int. J. Adv. Sci. Eng. Technol.*, 53–58.
- Javeed, A., Berglund, J. S., Dallora, A. L., Saleem, M. A., and Anderberg, P. (2023b). Predictive power of XGBoost_BiLSTM model: a machine-learning approach for accurate sleep apnea detection using electronic health data. *Int. J. Comput. Intell. Syst.* 16, 188. doi:10.1007/s44196-023-00362-y
- Javeed, A., Dallora, A. L., Berglund, J. S., Ali, A., Ali, L., and Anderberg, P. (2023c). Machine learning for dementia prediction: a systematic review and future research directions. *J. Med. Syst.* 47, 17. doi:10.1007/s10916-023-01906-7
- Javeed, A., Dallora, A. L., Berglund, J. S., Ali, A., Anderberg, P., and Ali, L. (2023d). Predicting dementia risk factors based on feature selection and neural networks. *Comput. Mater. Continua* 75, 2491–2508. doi:10.32604/cmc.2023.033783
- Javeed, A., Dallora, A. L., Berglund, J. S., and Anderberg, P. (2022b). An intelligent learning system for unbiased prediction of dementia based on autoencoder and adaboost ensemble learning. *Life* 12, 1097. doi:10.3390/life12071097
- Javeed, A., Dallora, A. L., Berglund, J. S., Idrisoglu, A., Ali, L., Rauf, H. T., et al. (2023e). Early prediction of dementia using feature extraction battery (feb) and optimized support vector machine (svm) for classification. *Biomedicine* 11, 439. doi:10.3390/biomedicine11020439
- Javeed, A., Rizvi, S. S., Zhou, S., Riaz, R., Khan, S. U., and Kwon, S. J. (2020). Heart risk failure prediction using a novel feature selection method for feature refinement and neural network for classification. *Mob. Inf. Syst.* 2020, 1–11. doi:10.1155/2020/8843115
- Javeed, A., Saleem, M. A., Dallora, A. L., Ali, L., Berglund, J. S., and Anderberg, P. (2023f). Decision support system for predicting mortality in cardiac patients based on machine learning. *Appl. Sci.* 13, 5188. doi:10.3390/app13085188
- Javeed, A., Zhou, S., Yongjian, L., Qasim, I., Noor, A., and Nour, R. (2019). An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection. *IEEE Access* 7, 180235–180243. doi:10.1109/access.2019.2952107
- Karaglanli, M., Gourlia, K., Tsamardinos, I., and Chatzaki, E. (2020). Accurate blood-based diagnostic biosignatures for alzheimer's disease via automated machine learning. *J. Clin. Med.* 9, 3016. doi:10.3390/jcm9093016
- Kühnel, L., Bouteloup, V., Lespinasse, J., Chêne, G., Dufouil, C., Molinuevo, J. L., et al. (2021). Personalized prediction of progression in pre-dementia patients based on individual biomarker profile: a development and validation study. *Alzheimer's Dementia* 17, 1938–1949. doi:10.1002/alz.12363
- Lagergren, M., Fratiglioni, L., Hallberg, I. R., Berglund, J., Elmståhl, S., Hagberg, B., et al. (2004). A longitudinal study integrating population, care and social services data. the Swedish national study on aging and care (snac). *Aging Clin. Exp. Res.* 16, 158–168. doi:10.1007/bf03324546
- Liu, C., Li, R., Wu, S., Che, H., Jiang, D., Yu, Z., et al. (2023a). Self-guided partial graph propagation for incomplete multiview clustering. *IEEE Trans. Neural Netw. Learn. Syst.* 1–14. doi:10.1109/tnnls.2023.3244021
- Liu, C., Wu, S., Li, R., Jiang, D., and Wong, H.-S. (2023b). Self-supervised graph completion for incomplete multi-view clustering. *IEEE Trans. Knowl. Data Eng.* 35, 9394–9406. doi:10.1109/tkde.2023.3238416
- Liu, H., and Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Trans. Knowl. Data Eng.* 17, 491–502. doi:10.1109/tkde.2005.66
- Liu, M., Lang, R., and Cao, Y. (2015). Number of trees in random forest. *Comput. Eng. Appl.* 51, 126–131. doi:10.1007/978-3-642-31537-4_13
- Maito, M. A., Santamaría-García, H., Moguilner, S., Possin, K. L., Godoy, M. E., Avila-Funes, J. A., et al. (2023). Classification of alzheimer's disease and frontotemporal dementia using routine clinical and cognitive measures across multicentric underrepresented samples: a cross sectional observational study. *Lancet Regional Health—Americas* 17, 100387. doi:10.1016/j.lana.2022.100387
- Moslehi, S., Rabiei, N., Soltanian, A. R., and Mamani, M. (2022). Application of machine learning models based on decision trees in classifying the factors affecting mortality of covid-19 patients in hamadan, Iran. *BMC Med. Inf. Decis. Mak.* 22, 192. doi:10.1186/s12911-022-01939-x
- Naseem, U., Rashid, J., Ali, L., Kim, J., Haq, Q. E. U., Awan, M. J., et al. (2022). An automatic detection of breast cancer diagnosis and prognosis based on machine learning using ensemble of classifiers. *IEEE Access* 10, 78242–78252. doi:10.1109/access.2022.3174599
- Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., Chalek, J., Abd-Allah, F., et al. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the global burden of disease study 2019. *Lancet Public Health* 7, e105–e125. doi:10.1016/s2468-2667(21)00249-8
- Nyholm, J., Ghazi, A. N., Ghazi, S. N., and Sanmartin Berglund, J. (2023). Prediction of dementia based on older adults' sleep disturbances using machine learning. Karlskrona, Sweden: Blekinge Institute of Technology.
- Patel, T., Polikar, R., Davatzikos, C., and Clark, C. M. (2008). "Eeg and mri data fusion for early diagnosis of alzheimer's disease," in 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, Canada, August, 2008, 1757–1760.
- Patnode, C. D., Perdue, L. A., Rossom, R. C., Rushkin, M. C., Redmond, N., Thomas, R. G., et al. (2020). Screening for cognitive impairment in older adults: updated evidence report and systematic review for the us preventive services task force. *Jama* 323, 764–785. doi:10.1001/jama.2019.22258
- Paul, A., Mukherjee, D. P., Das, P., Gangopadhyay, A., Chintia, A. R., and Kundu, S. (2018). Improved random forest for classification. *IEEE Trans. Image Process.* 27, 4012–4024. doi:10.1109/tip.2018.2834830
- Rostamzadeh, A., Schwegler, C., Gil-Navarro, S., Rosende-Roca, M., Romotzky, V., Ortega, G., et al. (2021). Biomarker-based risk prediction of alzheimer's disease dementia in mild cognitive impairment: psychosocial, ethical, and legal aspects. *J. Alzheimer's Dis.* 80, 601–617. doi:10.3233/jad-200484
- Ryu, S.-E., Shin, D.-H., and Chung, K. (2020). Prediction model of dementia risk based on xgboost using derived variable extraction and hyper parameter optimization. *IEEE Access* 8, 177708–177720. doi:10.1109/access.2020.3025553
- Ryzhikova, E., Ralbovsky, N. M., Sikirzhitsky, V., Kazakov, O., Halamkova, L., Quinn, J., et al. (2021). Raman spectroscopy and machine learning for biomedical applications: alzheimer's disease diagnosis based on the analysis of cerebrospinal fluid. *Spectrochimica Acta Part A Mol. Biomol. Spectrosc.* 248, 119188. doi:10.1016/j.saa.2020.119188
- Saleem, M. A., Thien Le, N., Asdornwiset, W., Chaitusaney, S., Javeed, A., and Benjapolakul, W. (2023). Sooty tern optimization algorithm-based deep learning model for diagnosing nslc tumours. *Sensors* 23, 2147. doi:10.3390/s23042147
- Salem, F. A., Chaaya, M., Ghannam, H., Al Feel, R. E., and El Asmar, K. (2021). Regression based machine learning model for dementia diagnosis in a community setting. *Alzheimer's Dementia* 17, e053839. doi:10.1002/alz.053839
- Salihović, D., Smajlović, D., Mijajlović, M., Zoletić, E., and Ibrahimagić, O. Č. (2018). Cognitive syndromes after the first stroke. *Neurol. Sci.* 39, 1445–1451. doi:10.1007/s10072-018-3447-6
- Salzberg, S. L. (1994). *C4.5: programs for machine learning by j. ross quinlan*. Cambridge, Massachusetts, United States: Morgan Kaufmann Publishers, Inc.
- Shahzad, A., Dadlani, A., Lee, H., and Kim, K. (2022). Automated prescreening of mild cognitive impairment using shank-mounted inertial sensors based gait biomarkers. *IEEE Access* 10, 15835–15844. doi:10.1109/access.2022.3149100
- Shigemizu, D., Akiyama, S., Asanomi, Y., Boroevich, K. A., Sharma, A., Tsunoda, T., et al. (2019). A comparison of machine learning classifiers for dementia with lewy bodies using mirna expression data. *BMC Med. Genomics* 12, 150–210. doi:10.1186/s12920-019-0607-3
- Sivakani, R., and Ansari, G. A. (2020). "Machine learning framework for implementing alzheimer's disease," in 2020 International conference on communication and signal processing (ICCCSP), Chennai, India, July, 2020, 0588–0592.
- Spooner, A., Chen, E., Sowmya, A., Sachdev, P., Kochan, N. A., Trollor, J., et al. (2020). A comparison of machine learning methods for survival analysis of high-dimensional clinical data for dementia prediction. *Sci. Rep.* 10, 20410. doi:10.1038/s41598-020-77220-w
- Stamate, D., Alghamdi, W., Ogg, J., Hoile, R., and Murtagh, F. (2018). "A machine learning framework for predicting dementia and mild cognitive impairment," in 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, USA, December, 2018, 671–678.

- Studholme, C., Cardenas, V., Song, E., Ezekiel, F., Maudsley, A., and Weiner, M. (2004). Accurate template-based correction of brain mri intensity distortion with application to dementia and aging. *IEEE Trans. Med. Imaging* 23, 99–110. doi:10.1109/tmi.2003.820029
- Tanveer, M., Richhariya, B., Khan, R. U., Rashid, A. H., Khanna, P., Prasad, M., et al. (2020). Machine learning techniques for the diagnosis of alzheimer's disease: a review. *ACM Trans. Multimedia Comput. Commun. Appl. (TOMM)* 16, 1–35. doi:10.1145/3344998
- Tuv, E., Borisov, A., Runger, G., and Torkkola, K. (2009). Feature selection with ensembles, artificial variables, and redundancy elimination. *J. Mach. Learn. Res.* 10, 1341–1366. doi:10.1145/1577069.1755828
- Visser, P. J., Lovestone, S., and Legido-Quigley, C. (2019). A metabolite-based machine learning approach to diagnose alzheimer-type dementia in blood: results from the european medical information framework for alzheimer disease biomarker discovery cohort. *Alzheimers Dement. (N Y)* 18. doi:10.1016/j.trci.2019.11.001
- Wang, J., Knol, M. J., Tiulpin, A., Dubost, F., de Bruijne, M., Vernooij, M. W., et al. (2019). Grey matter age prediction as a biomarker for risk of dementia: a population-based study. <https://www.biorxiv.org/content/10.1101/518506v1>.
- Yager, R. R. (2006). An extension of the naive bayesian classifier. *Inf. Sci.* 176, 577–588. doi:10.1016/j.ins.2004.12.006
- Yu, J.-T., Xu, W., Tan, C.-C., Andrieu, S., Suckling, J., Evangelou, E., et al. (2020). Evidence-based prevention of alzheimer's disease: systematic review and meta-analysis of 243 observational prospective studies and 153 randomised controlled trials. *J. Neurology, Neurosurg. Psychiatry* 91, 1201–1209. doi:10.1136/jnnp-2019-321913
- Zhang, Y., Zhang, H., Cai, J., and Yang, B. (2014). A weighted voting classifier based on differential evolution. *Abstr. Appl. analysis* 2014, 1–6. doi:10.1155/2014/376950
- Zheng, W., Zhu, X., Zhu, Y., Hu, R., and Lei, C. (2018). Dynamic graph learning for spectral feature selection. *Multimedia tools Appl.* 77, 29739–29755. doi:10.1007/s11042-017-5272-y