World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

# Classification and analysis of chronic health conditions influencing dementia progression using machine learning methods

Generaldo Maylem, [1,*], Isaac Angelo Dioses [2], Catleen Glo Feliciano [3], Jilian Dizon [4], Loida Hermosura CpE, MSIT [5], Bryan Tababa [3] and Zenus Labugen [3]

[1] *College of Medicine, Isabela State University, Echague Campus, Echague, Isabela, Philippines.*
[2] *School of Information Technology, MAPUA University, Makati Campus, Makati, Philippines.*
[3] *College of Computing Studies, Information and Communication Technology, Isabela State University, Echague Campus, Echague, Philippines.*
[4] *University of Lasalette College, Santiago City, Philippines.*
[5] *Department of Information Technology, Northeaster College, Santiago City, Philippines.*

World Journal of Advanced Research and Reviews, 2024, 24(03), 1600–1608

Publication history: Received on 07 November 2024; revised on 14 December 2024; accepted on 17 December 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.3.3848

## Abstract

Dementia, a progressive neurodegenerative condition, poses significant challenges to global healthcare systems due to its rising prevalence and complexity in diagnosis. This study employs a machine learning-based methodology to categorize dementia-related data and forecast disease progression, enabling early diagnosis and personalized treatment planning. Using a structured dataset comprising 24 features and 2230 entries, data preprocessing was conducted to address missing values, encode categorical variables, and normalize numerical features. Exploratory data analysis revealed key insights, including a strong inverse correlation between cognitive test scores and dementia severity. A Random Forest Classifier was trained to predict dementia presence, achieving an accuracy of 91.40%. Feature importance analysis identified "Cognitive_Test_Scores," "Prescription_Memantine," and "Dosage in mg" as the most critical predictors, aligning with clinical evidence. The confusion matrix highlighted high predictive accuracy but revealed misclassifications, particularly false negatives, which are critical in medical contexts. This study underscores the potential of machine learning in dementia diagnosis and management. By identifying significant predictors and achieving robust classification performance, the research provides a foundation for integrating artificial intelligence into healthcare systems to enhance diagnostic accuracy, optimize treatment plans, and improve patient outcomes. Further refinement and validation using larger datasets are recommended for broader applicability in clinical settings.

**Keywords:** Dementia; Correlation; Random Forest; Machine Learning; Medical AI

## 1. Introduction

In Western countries, dementia was associated with an estimated $604 billion in medical and care costs in 2010, making it a significant burden on healthcare systems and economies globally [1]. The scale of the issue continues to grow, with approximately ten million new cases of dementia diagnosed annually. Projections indicate that by 2050, the number of individuals living with dementia will reach a staggering 135 million worldwide, highlighting the urgency of addressing this global health challenge [2]. Age is the most prominent and well-established risk factor for dementia. The prevalence of dementia increases exponentially with advancing age, affecting approximately 1-2 percent of individuals at age 65. This prevalence rises sharply to 30 percent by the age of 85, emphasizing the critical role of age in dementia onset and progression. With global life expectancy increasing and populations aging, the prevalence of dementia is expected to rise substantially, placing additional strain on healthcare systems. Among the various types of dementia, Alzheimer's disease (AD) is the most common subtype, accounting for 60 to 90 percent of all neurodegenerative illnesses, depending

* Corresponding author: Generaldo Maylem

on the diagnostic criteria applied. AD is characterized by a progressive decline in cognitive function, memory, and the ability to perform everyday activities. Its classification as the predominant form of dementia underscores the need for targeted research, early diagnosis, and innovative treatment approaches to mitigate its widespread impact. Dementia significantly impacts families both financially and emotionally, as they face the challenges of caregiving, managing behavioral and psychological symptoms, and coping with the loss of their loved one's social and familial roles. The costs of care, including medical expenses, caregiving services, and lost income, add to the financial strain. Degenerative dementias like Alzheimer's disease, Parkinson's disease with dementia, and frontotemporal dementia account for most memory impairment cases, progressively worsening over time and requiring increasing care. These challenges highlight the need for better support systems, early intervention, and accessible resources for affected individuals and their families [3].

Recent advancements in diagnostic tools, such as amyloid and Tau positron emission tomography (PET) imaging, as well as the detection of cerebrospinal fluid and blood biomarkers including peptides, total tau, p181-tau, and neurofilament light chain, have markedly improved the accuracy of dementia diagnoses [4]. These innovations enable earlier and more precise identification of dementia subtypes, fostering better disease monitoring and tailored patient care. Despite these diagnostic strides, effective medical therapies to halt or significantly slow the progression of dementia remain elusive. Current treatments offer only modest symptomatic relief, leaving a significant unmet need for disease-modifying interventions [5]. The complexity of dementia, influenced by multifaceted factors such as aging, genetics, and comorbid chronic conditions, poses significant challenges to therapeutic development. Accurately predicting the course of dementia is a crucial tool for clinicians and medical policy decision-makers to guide both individual patient care and more public health initiatives [6]. Because of the intricate interactions between genetic, environmental, and lifestyle factors, the course of degenerative dementia, including Alzheimer's disease, Parkinson's disease with dementia, and frontotemporal dementia, varies greatly from person to person. It is difficult to create consistent treatment or care paths because of this heterogeneity, which highlights the need for predictive techniques that take individual characteristics into consideration. Advanced imaging methods and fluid-based biomarkers are examples of integrated biomarker profiles that have shown great promise in improving prediction accuracy. These biomarkers aid doctors in patient stratification and more accurate prediction of the rate of cognitive decline by offering insights into underlying disease mechanisms, such as neurodegeneration, tau pathology, and amyloid deposition. However, several obstacles prevent their widespread use in clinical practice. These devices' exorbitant cost restricts accessibility, especially in environments with little resources, and raises issues with healthcare equity. Additionally, the use of biomarkers presents ethical concerns about psychological discomfort, possible stigma, and the consequences of predictive knowledge for patients and their families, particularly in asymptomatic or early-stage individuals [7,8,9,10].

For physicians and families to provide patients with more individualized and efficient care, early detection signs that indicate the advancement of dementia are essential. Early detection enables prompt interventions that can improve quality of life, reduce the progression of the disease, and better prepare families for the difficulties that lie ahead. Dementia prediction models are now crucial instruments in clinical practice for assessing risk and development. These models combine lifestyle and demographic characteristics to provide a thorough method of evaluating personal vulnerabilities. In dementia risk profiling, demographic variables including age, education level, and body mass index (BMI) are important. The greatest known risk factor is getting older, and lower educational attainment is linked to a smaller cognitive reserve, which may hasten the onset of symptoms.

### 1.1. Machine Learning Methods

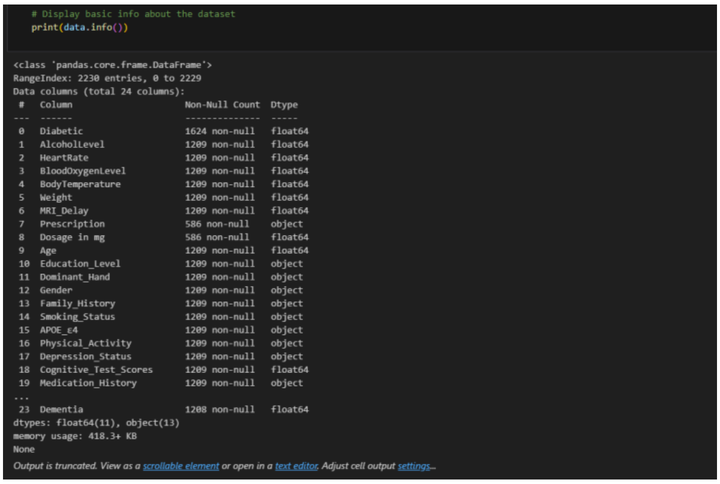
Medical imaging frequently uses machine learning (ML) methods, especially deep learning models like convolutional neural networks (CNNs), to identify anomalies like cancers in MRI scans and nodules in CT scans. According to studies, these models are more accurate and efficient than conventional diagnostic techniques, which makes them incredibly useful for physicians in the early detection of crucial conditions [13], [14]. By examining biomarkers and neuroimaging data, machine learning has also proven crucial in forecasting the course of diseases like Parkinson's and Alzheimer's [15]. ML makes it possible to analyze proteome, metabolomic, and genetic data in the field of personalized medicine in order to forecast treatment outcomes and adverse medication reactions. ML models, for instance, are used to determine patient-specific medication schedules, which lowers the possibility of ineffective treatments and speeds up the drug discovery process [16]. These developments demonstrate how ML can be used to develop customized healthcare solutions that increase the effectiveness of treatment [17]. DCNNs have been effectively utilized to detect diabetic retinopathy (DR) from retinal fundus images. Studies have demonstrated that these networks can classify the severity of DR with high accuracy, facilitating early diagnosis and timely treatment to prevent vision loss. For instance, a study employed a CNN architecture to classify DR severity using the EyePACS dataset, achieving promising results in automating DR diagnosis [18]. The early diagnosis and detection of Alzheimer's disease (AD) is one of the most well-known uses of machine learning. To find indicators like amyloid-beta deposits and hippocampus shrinkage, ML models in particular, deep learning frameworks like convolutional neural networks (CNNs) have proved crucial in evaluating neuroimaging data, such as structural MRI and PET scans. Research utilizing the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset shows that machine learning algorithms are highly accurate at differentiating between healthy individuals, mild cognitive impairment (MCI), and Alzheimer's disease [19]. Additionally, ML is essential for the identification and categorization of brain tumors. Brain tumors may be accurately identified and segmented from MRI and CT scans using advanced machine learning models, such as CNNs and topologies like U-Net. This allows for the precise delineation of tumor boundaries for surgical planning. When it comes to classifying tumor subregions such edema, necrosis, and enhancing tumor tissues, these models routinely rank among the best in the Brain Tumor Segmentation (BraTS) tasks [20]. By enhancing decision-making and treatment approaches, reinforcement learning techniques have further optimized processes in the diagnosis of brain tumors .

## 2. Material and methods

To categorize dementia-related data and forecast the course of the disease, this study uses a machine learning-based methodology. Accurate classification models can help doctors with early diagnosis, individualized treatment planning, and better patient care as dementia prevalence rises. To process structured datasets, train models to find significant patterns, and assess their forecast performance using reliable metrics, the technique was created especially for this purpose. Data preprocessing, feature selection, model training, assessment, and validation are the main steps in the workflow, and each is essential to obtaining high model accuracy and dependability.

### 2.1. Data Analysis

The dataset used in this study contains 24 features and 2230 entries, consisting of a mix of numerical and categorical variables. However, it is observed that some columns have missing values, which need to be addressed to ensure the data is suitable for machine learning applications. The preprocessing process begins with handling these missing values, as several columns, such as "Diabetic," "AlcoholLevel," and "Cognitive_Test_Scores," have incomplete data. Numerical columns with missing values are imputed using statistical measures like the mean or median, ensuring that the overall distribution of the data is preserved. For categorical columns, missing values are replaced with the most frequent category or labeled as "Unknown" when appropriate. Columns with significant missingness, such as "Prescription" and "Dosage in mg," are carefully evaluated to determine whether they should be retained or excluded based on their importance to the classification task. To ensure consistency, numerical columns are standardized or normalized, bringing all features to the same scale. This step is essential to prevent features with larger ranges from dominating the training process. For categorical variables, the dataset is prepared by encoding these variables into numerical formats to make them compatible with machine learning models. Ordinal variables, such as "Education_Level," are encoded using techniques that preserve the order of the categories, while nominal variables, like "Gender" or "Smoking_Status," are transformed using one-hot encoding to create binary representations in Figure 1. Shows the pre-processing phase.



```
# Display basic info about the dataset
print(data.info())

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2230 entries, 0 to 2229
Data columns (total 24 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   Diabetic               1624 non-null   float64
 1   AlcoholLevel           1209 non-null   float64
 2   HeartRate              1209 non-null   float64
 3   BloodOxygenLevel       1209 non-null   float64
 4   BodyTemperature        1209 non-null   float64
 5   Weight                 1209 non-null   float64
 6   MRI_Delay              1209 non-null   float64
 7   Prescription           586 non-null    object
 8   Dosage in mg           586 non-null    float64
 9   Age                    1209 non-null   float64
 10  Education_Level        1209 non-null   object
 11  Dominant_Hand          1209 non-null   object
 12  Gender                 1209 non-null   object
 13  Family_History         1209 non-null   object
 14  Smoking_Status         1209 non-null   object
 15  APOE_ε4                1209 non-null   object
 16  Physical_Activity      1209 non-null   object
 17  Depression_Status      1209 non-null   object
 18  Cognitive_Test_Scores  1209 non-null   float64
 19  Medication_History     1209 non-null   object
...
 23  Dementia               1208 non-null   float64
dtypes: float64(11), object(13)
memory usage: 418.3+ KB
None
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

**Figure 1** Data Preprocessing

This bar chart in figure 2 illustrates the relationship between different chronic health conditions diabetes, heart disease, and hypertension and their proportions among dementia cases. The y-axis represents the proportion of dementia cases, while the x-axis categorizes the chronic health conditions. The data reveals that individuals with diabetes have the highest proportion of dementia cases, followed closely by those with hypertension and heart disease. This suggests a potential link between these chronic conditions and an increased likelihood of developing dementia.

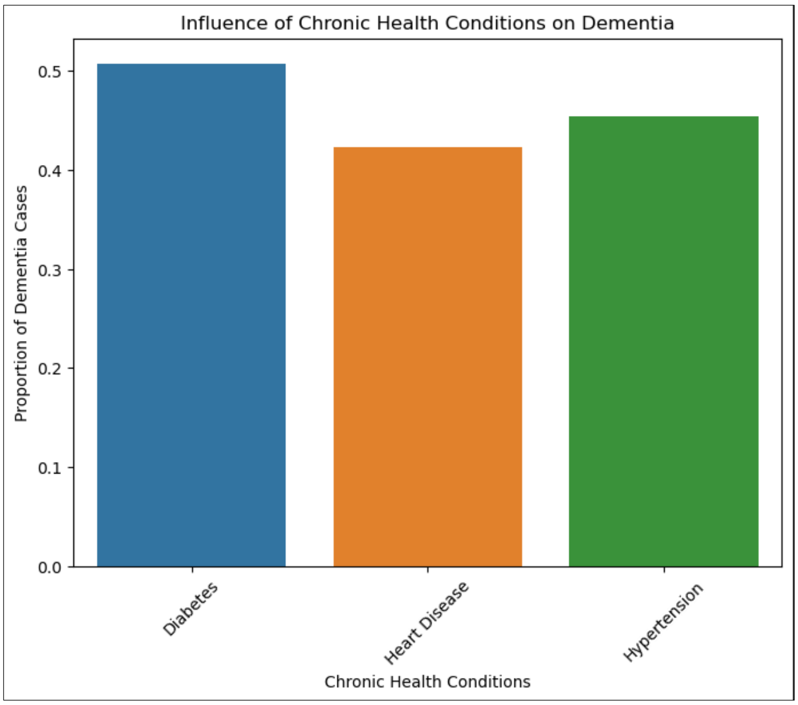



**Figure 2** Influence of Chronic Health Conditions on Dementia

In the Figure 3. illustrates the distribution of individuals across different age groups, categorized by dementia status (with 0.0 indicating no dementia and 1.0 indicating dementia). The y-axis represents the count of individuals, while the x-axis represents age ranges. Stacked bars are used to show the breakdown of dementia presence within each age group, and KDE (kernel density estimation) lines provide an additional visualization of the age distribution trends. The data reveals that dementia cases (orange bars) are more prevalent in the younger elderly age group around 60–65 years and then again at older ages closer to 90. Meanwhile, individuals without dementia (blue bars) dominate most age groups in the mid-range ages, especially around 70–80 years. The KDE lines further highlight these trends, with the blue curve peaking in the mid-70s and the orange curve peaking around the early 60s.

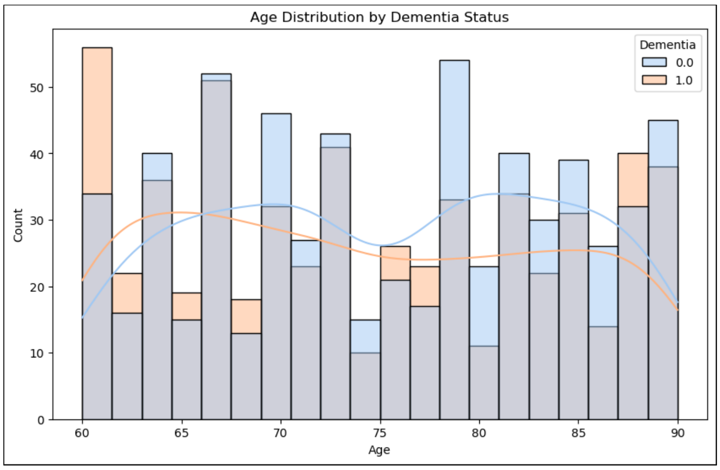



**Figure 3** Age Distribution by Dementia Status

## 2.2. Descriptive Statistics

Descriptive statistics are generated to summarize the dataset. For numeric features like "BloodOxygenLevel," "BodyTemperature," and "Cognitive_Test_Scores," measures such as mean, standard deviation, and range are calculated to understand their distribution. Categorical features, including "Gender," "Family_History," and "Physical_Activity," are analyzed through frequency counts to determine the prevalence of different categories.

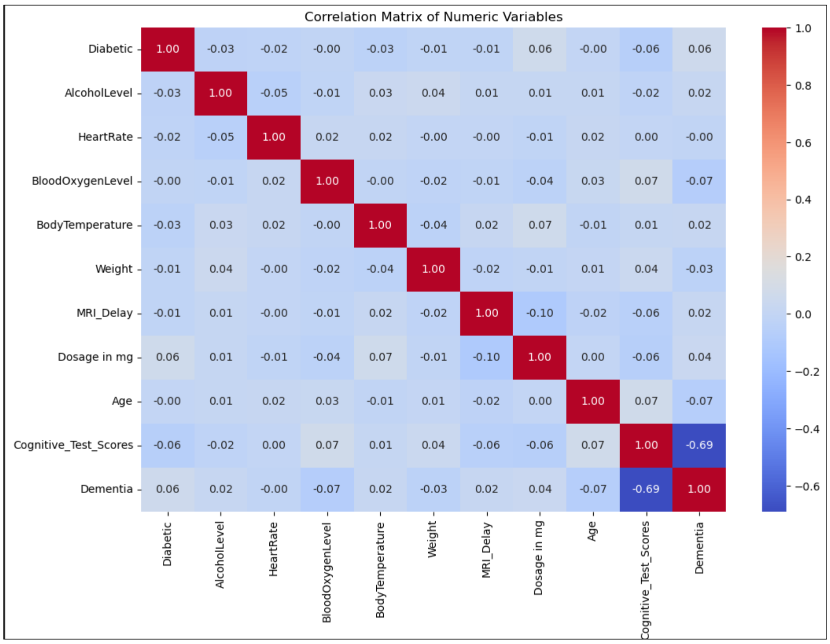



**Figure 4** Correlation Matrix of Numeric Variables Between Features

In the dataset, the strongest correlation is observed between cognitive test scores and dementia, with a coefficient of -0.69. This negative correlation suggests that lower cognitive test scores are strongly associated with the presence or severity of dementia, indicating a meaningful inverse relationship. Such a correlation aligns with the clinical understanding that declining cognitive abilities are a key indicator of dementia progression. Another notable, albeit weaker, correlation is between age and cognitive test scores, which shows a slightly positive trend. However, this relationship is not statistically significant and may be influenced by other variables. The analysis further reveals that most correlations among the other variables, such as blood oxygen levels, weight, or body temperature, are weak and do not indicate substantial linear relationships. These findings highlight the importance of focusing on the strongest correlations, such as the one between cognitive test scores and dementia, as they provide valuable insights for modeling and understanding the underlying factors contributing to dementia in the studied population. By emphasizing these relationships, the methodology ensures a data-driven approach to identifying significant predictors and patterns within the dataset.

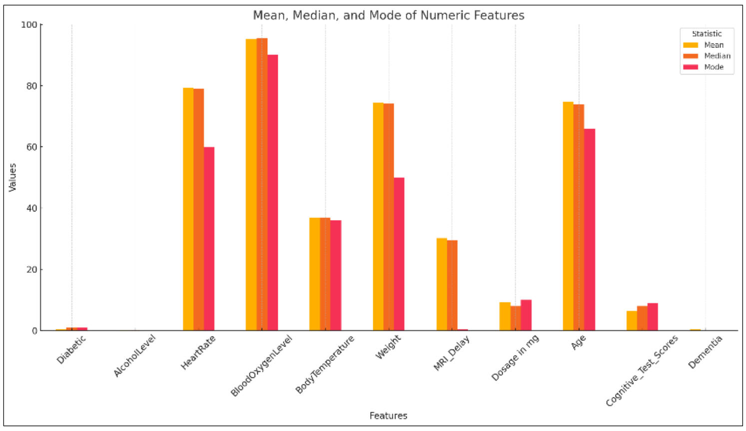



**Figure 5** Mean Median Mode of Numeric Features

In Figure 5 shows additional descriptive statistics. The "HeartRate" feature has a mean of 79.32 and a median of 79, indicating a relatively symmetric distribution with no significant outliers skewing the central values. Similarly, "BloodOxygenLevel" has a mean of 95.23 and a median of 95.43, showing a highly concentrated range of values close to the average. On the other hand, the "Dosage in mg" feature indicates a minor tilt in the distribution toward higher values, with a mean of 9.16 and a median of 8. Although the median stays closer to the data's middle cluster, this disparity implies that some people are receiving noticeably greater dosages, which is pushing the mean upward. The median for "Cognitive_Test_Scores," the mode, and the mean are 8.36, 8.36, and 9. A clustering of values around 9 is reflected by the mode being higher than the mean and median, but the presence of lower scores lowers the mean and results in a distribution that is left-skewed. This suggests that the distribution is impacted by a subset of people with extremely low cognitive scores. In the "Dementia" feature, which is binary, the mode is 0, reflecting the dominance of non-dementia cases. The mean of 0.48 aligns with the binary nature of the variable, suggesting an almost equal proportion of cases and non-cases, though non-dementia cases slightly outnumber dementia cases. Lastly, "MRI_Delay" has a mean of 30.14 and a median of 29.53, with a mode of 0.34. The notable gap between the mode and the mean suggests that most observations are clustered at very low delay times, while a subset of higher delay times elevates the mean. This highlights a distribution with a concentration of lower values but a tail extending to higher values.

### 2.3. Machine Learning Process

After analysis and visualization, researcher proceed to Machine Learning for classification and prediction of Dementia based on features. The dataset is divided into features (X) and the target variable (y). The features consist of all columns except for the "Dementia" column, which serves as the target. This separation ensures that the model is trained to predict the presence of dementia based on the other variables. The data is then split into training and testing sets, allocating 75% of the data for training and 25% for testing. The train_test_split function is used with a fixed random state to ensure reproducibility. A Random Forest Classifier is employed to train the model. This ensemble learning method builds multiple decision trees and aggregates their outputs to improve prediction accuracy and reduce overfitting. The model is trained using the training dataset (X_train and y_train), learning patterns that can distinguish between dementia and non-dementia cases. Once the model is trained, feature importance is calculated using the feature_importances_ attribute of the Random Forest Classifier. This identifies how much each feature contributes to the model's predictions. The importance scores are organized into a DataFrame, which ranks features in descending order of importance. This process helps to identify which variables are the most significant predictors of dementia in the dataset.

```
data['Dementia'] = data['Dementia'].astype(int)
categorical_columns = data.select_dtypes(include=['object']).columns
```

```
data_encoded = pd.get_dummies(data, columns=categorical_columns, drop_first=True)

X = data_encoded.drop('Dementia', axis=1)
y = data_encoded['Dementia']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.25, random_state=42)
```

```
model = RandomForestClassifier(random_state=42)
model.fit(X_train, y_train)
```

**Figure 6** Machine Learning Code

## 3. Results and Discussion

The importance score for each feature reflects its relative weight in determining the model's outcome. Features with higher scores have a stronger impact on the model's predictions, while those with lower scores contribute less in figure 7 shows the feature importance ranking. The most significant predictor in the dataset is Cognitive_Test_Scores, which has the highest importance score. This suggests that cognitive test performance plays a critical role in distinguishing between dementia and non-dementia cases. Its dominance aligns with clinical evidence, as cognitive decline is a hallmark symptom of dementia. The second most important feature is Prescription_Memantine, which represents whether a patient is prescribed Memantine, a medication commonly used to treat symptoms of Alzheimer's disease. Its high importance score indicates that this prescription is strongly associated with dementia cases, likely due to its targeted use in dementia management. Dosage in mg is the third most influential feature, reflecting the dosage of medications prescribed to the patients. This indicates that higher or specific medication dosages may be predictive of dementia, potentially pointing to the severity or progression of the condition. Depression_Status_Yes also has a notable importance score, showing a connection between depression and dementia. This relationship is well-documented, as depression is often a comorbidity or an early sign of cognitive impairment. Other features like BloodOxygenLevel, Weight, and AlcoholLevel have moderate contributions, suggesting they have some predictive value but are less influential compared to the cognitive and prescription-related variables. Features like BodyTemperature and MRI_Delay have the lowest importance scores, indicating minimal impact on the prediction of dementia in this dataset.

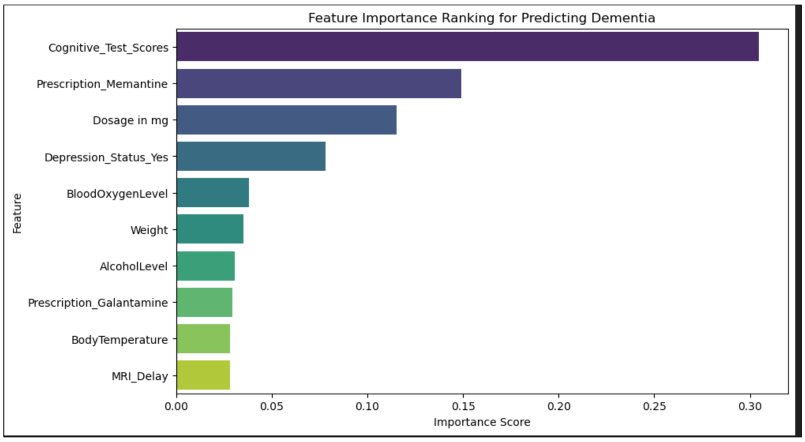



**Figure 7** Feature Importance ranking

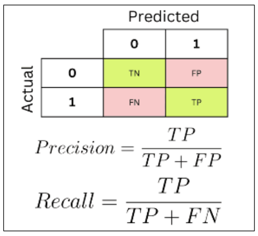



**Figure 8** Confusion Matrix Formula

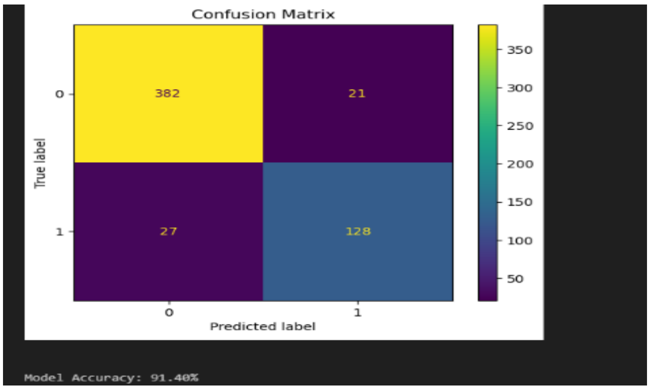



**Figure 9** Confusion Matrix Results

The confusion matrix represents the performance of the dementia prediction model by displaying the actual versus predicted classifications. The top-left value (382) indicates the number of true negatives, where the model correctly predicted individuals without dementia. The bottom-right value (128) represents the true positives, where the model accurately predicted dementia cases. The top-right value (21) corresponds to false positives, where the model incorrectly classified individuals without dementia as having dementia. Conversely, the bottom-left value (27) represents false negatives, where the model failed to identify dementia cases and classified them as non-dementia. The model's accuracy is reported as 91.40%, which is calculated as the ratio of correctly predicted cases (both true positives and true negatives) to the total number of predictions. While the overall accuracy is high, the confusion matrix highlights that some false negatives and false positives remain, which could have implications for the reliability of the model in practical scenarios. False negatives are critical in medical contexts as they represent undetected dementia cases, which might lead to delayed diagnosis and treatment. This analysis suggests that while the model is generally effective, further refinement might be necessary to reduce misclassification, especially false negatives, to improve clinical reliability.

## 4. Conclusion

This study demonstrates the application of a machine learning-based methodology to analyze and predict dementia-related outcomes. By preprocessing the dataset with 24 features and addressing missing values, the data was transformed into a format suitable for model training and evaluation. Exploratory data analysis revealed meaningful insights, such as the strong inverse correlation between cognitive test scores and dementia, highlighting the significance of cognitive performance in the diagnosis and progression of dementia. Other features, such as the use of prescription medications like Memantine, dosage in milligrams, and depression status, also emerged as critical predictors, further validating their clinical relevance.

The Random Forest Classifier achieved a high accuracy of 91.40%, demonstrating its capability to effectively distinguish between dementia and non-dementia cases. Feature importance rankings identified "Cognitive_Test_Scores" as the most influential variable, aligning with existing clinical knowledge. Despite the high accuracy, the confusion matrix analysis revealed some misclassifications, particularly false negatives, which are critical in medical applications as they could result in undiagnosed cases. These findings suggest that further optimization, such as hyperparameter tuning or the inclusion of additional data, could enhance the model's performance and reliability.

## Compliance with ethical standards

### *Disclosure of Conflict of interest*

The authors declare that there is no conflict of interest related to the publication of this manuscript.

## References

[1] Prince MJ, Wimo A, Guerchet MM, Ali GC, Wu Y-T, Prina M. World Alzheimer Report 2015 - The Global Impact of Dementia: An analysis of prevalence, incidence, cost and trends. London: Alzheimer's Disease International; 2015. p. 84.

[2] 1.World Health Organization. Dementia [Internet]. Who.int. World Health Organization: WHO; 2019. Available from: https://www.who.int/health-topics/dementia#tab=tab_1

[3] Goodman RA, Lochner KA, Thambisetty M, Wingo TS, Posner SF, Ling SM. Prevalence of dementia subtypes in United States Medicare fee-for-service beneficiaries, 2011-2013. Alzheimer's & Dementia. 2016 May 9;13(1):28–37.

[4] Fink HA, Linskens EJ, Silverman PC, McCarten JR, Hemmy LS, Ouellette JM, et al. Accuracy of Biomarker Testing for Neuropathologically Defined Alzheimer Disease in Older Adults With Dementia. Annals of Internal Medicine. 2020 May 19;172(10):669–77.

[5] Cummings J, Lee G, Nahed P, Kambar MEZN, Zhong K, Fonseca J, et al. Alzheimer's disease drug development pipeline: 2022. Alzheimer's & Dementia: Translational Research & Clinical Interventions. 2022 Jan;8(1).

[6] Ren L, Liang J, Wan F, Wang Y, Dai X. Development of a Clinical Risk Score Prediction Tool for 5-, 9-, and 13-Year Risk of Dementia. JAMA Network Open. 2022 Nov 17;5(11):e2242596.

[7] Ranjan Duara, Barker WW. Heterogeneity in Alzheimer's Disease Diagnosis and Progression Rates: Implications for Therapeutic Trials. Neurotherapeutics. 2022 Jan 27;19(1):8–25.

[8] Jack CR, J. Scott Andrews, Beach TG, Buracchio T, Dunn B, Graf A, et al. Revised criteria for diagnosis and staging of Alzheimer's disease: Alzheimer's Association Workgroup. Alzheimer's & dementia. 2024 Jun 27;20(8).

[9] Kühnel L, Bouteloup V, Lespinasse J, Chêne G, Dufouil C, Molinuevo JL, et al. Personalized prediction of progression in pre-dementia patients based on individual biomarker profile: A development and validation study. Alzheimer's & Dementia. 2021 Sep 28;17(12):1938–49.

[10] Vanderschaeghe G, Dierickx K, Vandenberghe R. Review of the Ethical Issues of a Biomarker-Based Diagnoses in the Early Stage of Alzheimer's Disease. Journal of Bioethical Inquiry [Internet]. 2018 Mar 12;15(2):219–30. Available from: https://link.springer.com/article/10.1007/s11673-018-9844-y

[11] Kivipelto M, Ngandu T, Laatikainen T, Winblad B, Soininen H, Tuomilehto J. Risk score for the prediction of dementia risk in 20 years among middle aged people: a longitudinal, population-based study. The Lancet Neurology. 2006 Sep;5(9):735–41.

[12] Barnes DE, Beiser AS, Lee A, Langa KM, Koyama A, Preis SR, et al. Development and validation of a brief dementia screening indicator for primary care. Alzheimer's & Dementia. 2014 Feb 3;10(6):656.

[13] Huang YC, Liu TC, Lu CJ. Establishing a machine learning dementia progression prediction model with multiple integrated data. BMC Medical Research Methodology. 2024 Nov 22;24(1).

[14] Jia H, Zhang J, Ma K, Qiao X, Ren L, Shi X. Application of convolutional neural networks in medical images: a bibliometric analysis. Quantitative Imaging in Medicine and Surgery. 2024 May 1;14(5):3501–18.

[15] Lee JG, Jun S, Cho YW, Lee H, Kim GB, Seo JB, et al. Deep Learning in Medical Imaging: General Overview. Korean Journal of Radiology. 2017;18(4):570.

[16] Paul D, Sanap G, Shenoy S, Kalyane D, Kalia K, Tekade RK. Artificial intelligence in drug discovery and development. Drug Discovery Today [Internet]. 2020 Oct;26(1):80–93. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7577280/

[17] None Deepak Kumar, Pawar P, None Hari Gonaygunta, Nadella S, None Karthik Meduri, None Shoumya Singh. Machine learning's role in personalized medicine & treatment optimization. World Journal Of Advanced Research and Reviews. 2024 Feb 28;21(2):1675–86.

[18] Suedumrong C, Phongmoo S, Akarajaka T, Leksakul K. Diabetic Retinopathy Detection Using Convolutional Neural Networks with Background Removal, and Data Augmentation. Applied Sciences. 2024 Sep 30;14(19):8823.

[19] ADNI | Alzheimer's Disease Neuroimaging Initiative [Internet]. ADNI. Available from: https://adni.loni.usc.edu/

[20] Abdusalomov AB, Mukhiddinov M, Whangbo TK. Brain Tumor Detection Based on Deep Learning Approaches and Magnetic Resonance Imaging. Cancers. 2023 Aug 18;15(16):4172–2.