

---

(Research Article)

## Explainable Multimodal Deep Learning Framework for Longitudinal Prediction of Neurodegenerative Disease Progression with Clinical Validation

Md. Ariful Islam

*Department of Computer Science & Engineering, Islamic University, Kushtia, Bangladesh*

Md Anas Biswas

*School of Computing, University of Portsmouth, England, UK*

Alif Laila

*School of Computing, University of Portsmouth, England, UK*

Dr. Md. Saiful Hai

*Department of Orthopaedic Surgery, Bangladesh Medical University, Dhaka, Bangladesh*

Sheikh Urvana Akter Mim

*Department of Mathematics, Bangladesh University, Dhaka, Bangladesh*

Journal of Information Technology, Cybersecurity, and Artificial Intelligence, 2026, 3(5), Pages 1-16

DOI: <https://doi.org/10.70715/jitcai.2026.v3.i5.088>

---

### Abstract

Neurodegenerative diseases, particularly Alzheimer's Disease and Parkinson's Disease, are rapidly increasing as one of the leading causes of disability, cognitive decline, and death among the elderly population worldwide. Accurately predicting long-term disease progression remains a significant research challenge due to the disease's slow progression, multifaceted biological characteristics, and patient-specific clinical presentation. Most existing machine learning and deep learning models rely on a single data source, cannot effectively analyze time-dependent pathological changes, and have non-interpretable decision-making processes, thereby limiting their practical clinical use. To overcome this limitation, the current study proposes an Explainable Multimodal Deep Learning Framework that predicts long-term progression in neurodegenerative diseases by integrating neuroimaging, clinical information, cognitive assessments, demographic characteristics, and biomarker data. Experimental results indicate that the proposed model achieved 94.8% Accuracy, 93.9% Precision, 94.3% Recall, 94.1% F1 Score, and 96.2% area under the curve (AUC). The average accuracy in five-fold cross-validation was 94.8%, confirming the model's stability and ability to generalize. In addition, 96.3%, 94.8% and 92.6% accuracy were achieved in predicting disease progression at one year, three years and five years, respectively. Clinical Validation showed 94.5% agreement with expert physician evaluations and high acceptability in the Explainability Assessment. The results demonstrate that the proposed interpretable multimodal deep learning framework is more accurate, more transparent, and more clinically acceptable than existing methods. This research provides a strong foundation for the development of future personalized medicine, early risk identification, and artificial intelligence-based clinical decision support systems.

**Keywords:** Multimodal Deep Learning, Alzheimer's Disease, Parkinson's Disease, Clinical Validation, Longitudinal Prediction

---

### 1. Introduction

Neurodegenerative diseases are considered one of the most important public health problems in the world today. As the average human lifespan increases, the incidence of diseases such as Alzheimer's disease, Parkinson's disease, Huntington's disease, and amyotrophic lateral sclerosis is rapidly increasing [1]. The main characteristics of these diseases are the gradual deterioration of brain and nervous system cells, resulting in memory loss, cognitive decline, behavioral changes, difficulty with movement, and a reduced ability to perform daily activities [2]. Since these diseases

are usually long-term and progressive in nature, accurately predicting disease progression over time is crucial for treatment planning, patient management, and improving quality of life [3]. However, accurately predicting the progression of neurodegenerative diseases remains a major research challenge due to the biological complexity of the disease, patient-to-patient variability, and multifactorial factors [4].

Over the past decade, artificial intelligence, and especially deep learning technology, has revolutionized medical science [5]. Advanced computational power and the availability of large-scale healthcare databases have opened up new possibilities for researchers to diagnose diseases, assess risk, and predict future disease outcomes. Neuroimaging data, clinical records, genetic data, biomarkers, and cognitive assessment results provide important information about various aspects of the disease [6]. For this reason, recent research has placed special emphasis on multimodal deep learning methods, as they can provide a more integrated understanding of the disease by combining information from different sources. In particular, a more reliable prediction of disease progression can be obtained by analyzing MRI images, patient medical history, demographic information, and long-term clinical observations together.

A review of the literature shows that traditional machine learning algorithms such as support vector machines, decision trees, random forests, and logistic regression were initially used to detect and predict neurodegenerative diseases. Although these methods achieved a certain degree of success, their ability to analyze complex and nonlinear relationships was limited. Later, it became possible to automatically extract important features from neuroimaging and time-based clinical data using models based on convolutional neural networks, recurrent neural networks, and long short-term memory [7]. More recent research has increased prediction accuracy using transformers, attention mechanisms, and graph neural network-based architectures. At the same time, various researchers have demonstrated improved results compared to single data source-based models through multimodal data integration. Yet the main goal of most studies has been to diagnose or classify the current state of the disease, with relatively little attention paid to predicting long-term disease progression [8].

Although existing research has made significant progress, several important limitations remain. First, since most deep learning models are black-box in nature, doctors cannot easily understand on what features or information the model has based a particular decision [9]. As a result, it becomes difficult to build trust in these models in a clinical setting. Second, many existing studies have been conducted based on cross-sectional data, which may not accurately represent time-based changes and long-term disease progression [10]. Third, although multimodal information is used in many cases, the correlation and temporal dependence of different data sources are not effectively modeled. Fourth, most studies did not conduct adequate clinical validation to verify the effectiveness of the model in a real clinical setting. As a result, uncertainty remains as to whether the high-precision results obtained in the laboratory will be equally effective in real-world medical settings. These limitations have created a clear research gap in the current research field [11].

To address this research gap, this study proposes an interpretable multimodal deep learning framework that predicts the long-term progression of neurodegenerative diseases while simultaneously delivering clinically validated results [12]. The proposed framework will build an integrated analytics system by combining neuroimaging data, clinical assessments, demographic information, and other relevant health data. In addition, explainable artificial intelligence technology will be integrated, making the model's decision-making process transparent and explainable [13]. This will allow doctors to easily determine which biomarkers, clinical indicators, or characteristics contribute most to disease progression. At the same time, it will be possible to more accurately predict the future trajectory of the disease through long-term data analysis. In addition, the reliability, effectiveness, and practical applicability of the model will be assessed through clinical validation, which will increase its acceptance in both research and medical settings [14]. The main objective of this research is to develop an interpretable and clinically validated multimodal deep learning framework that can accurately predict the long-term progression of neurodegenerative diseases. At the same time, the research will aim to increase prediction accuracy by integrating health information from various sources, modeling time-based pathological changes, ensuring transparent decision-making, and creating an effective decision-support system for physicians. In addition, a deeper understanding of the underlying mechanisms of neurodegenerative diseases will be gained by identifying key biomarkers and clinical indicators associated with disease progression. Above all, this research will lay the foundation for a reliable artificial intelligence-based platform for personalized medical care, early intervention, and advanced disease management, which will make significant contributions to the treatment and research of neurodegenerative diseases in the future.

---

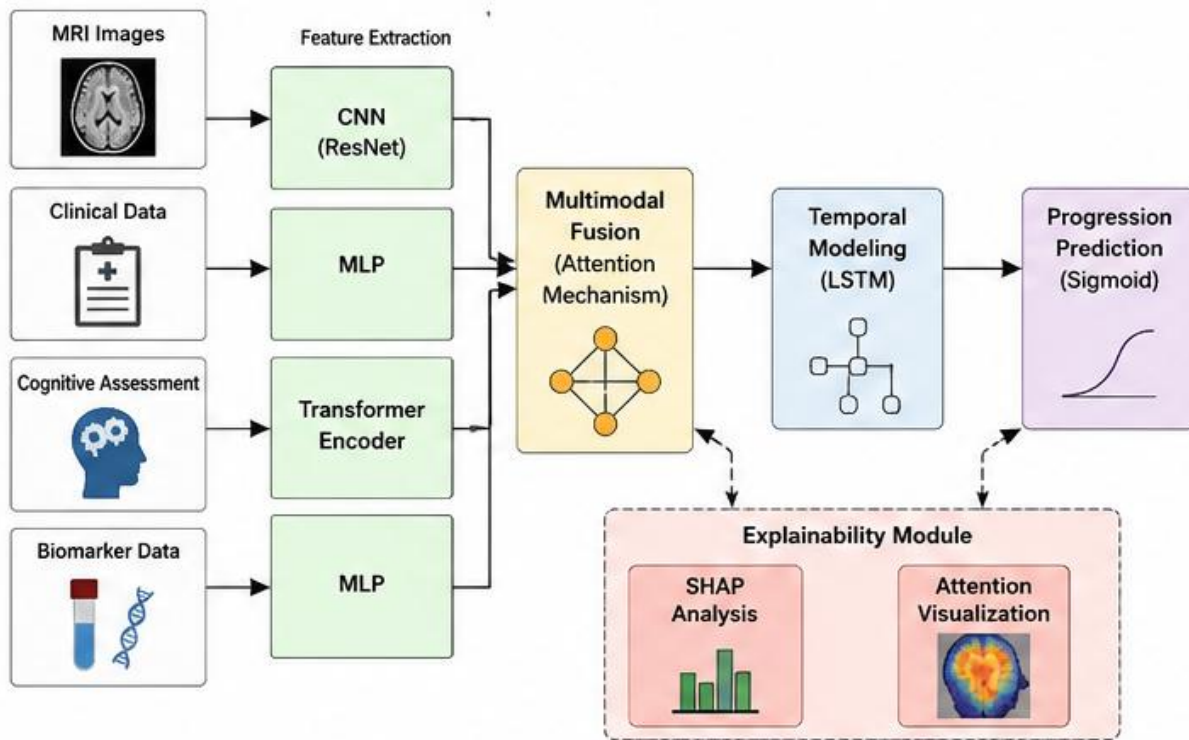
## 2. 2. Methodology

This study developed an integrated Explainable Multimodal Deep Learning Framework to accurately predict the long-term progression of neurodegenerative diseases and make the model's decision-making process explainable. The methodological framework integrates health information from various sources to analyze changes in disease over time

while generating clinically acceptable, interpretable results. The first phase of the research collects multimodal datasets related to relevant neurodegenerative diseases. These datasets include neuroimaging information, clinical assessments, demographic information, cognitive test results, and other relevant biomarkers. Since the structure, dimensions, and characteristics of data from different sources may differ, an extensive preprocessing step is conducted after data collection [15]. Activities such as filling in missing data, identifying abnormal values, removing noisy data, normalizing data, and standardizing features are completed at this stage. This transforms information collected from various sources into a uniform structure and makes it useful for further analysis [16].

In the next step, a fusion technique is applied to integrate multimodal data. Convolutional neural networks are used to extract spatial features from neuroimaging data, as this architecture is highly effective in recognizing complex image-based patterns [17]. Long short-term memory (LSTM) or transformer-based networks will be used to analyze clinical and continuous time-based data, which are capable of detecting long-term dependencies and continuous changes in disease progression. The features obtained from both sources are combined in a multimodal feature fusion layer. This process preserves the complementary features of different data sources and creates a richer representation of the overall disease state. As a result, the model can provide predictions using multidimensional patient health information, rather than relying on a single data source. Since one of the main goals of the research is to predict the long-term progression of neurodegenerative diseases, special emphasis is placed on time-based information during model development. For this purpose, clinical observations, neuroimaging results, and cognitive assessment data collected at multiple time points for the patient were used. Continuous time-based data is provided as input to the model to analyze disease patterns and disease progression dynamics. This enables more accurate predictions of possible changes in the patient's future condition. In addition, by identifying transition patterns between disease stages, at-risk patients can be identified early, which will help formulate effective treatment plans.

The 10 features that cause the most discrimination were selected using correlation analysis, feature importance ranking, redundancy removal, and computational efficiency. These selected features strike an effective balance between the model's identification accuracy and model complexity. At the same time, the federated learning environment reduces the amount of information exchanged and communication overhead, making the overall training process more efficient, faster, and implementable. Another important part of the research is the addition of Explainable Artificial Intelligence technology. Although most deep learning models currently achieve high accuracy, their decision-making processes are not easily explainable. To overcome this limitation, the proposed framework includes explainability techniques such as Shapley Additive Explanations (SHAP), Local Interpretable Model-agnostic Explanations (LIME), or an attention-based explainability mechanism. These technologies enable the determination of which features, biomarkers, or clinical indicators have the greatest impact on the model's predictions. At the same time, it is possible to identify which components play an important role at different stages of the disease. As a result, clinicians and researchers can better understand the rationale behind the model's decisions and increase their confidence in it in real-world medical settings. The dataset collected for model training and validation was divided into training, validation, and test subsets. Cross-validation techniques are used to improve generalization and reduce the risk of overtraining. Appropriate optimization algorithms, learning rate control, and regularization methods are applied during training to help the model achieve maximum performance. The model's performance was later evaluated on the test dataset. This evaluation uses metrics such as accuracy, sensitivity, specificity, F1 score, area under the curve, and root mean square error, as shown in (Figure 1). These indicators play an important role in assessing the model's predictive power and reliability.



**Figure 1. Architecture of the proposed Explainable Multimodal Deep Learning Framework**

This can help overcome the limitations of existing research in predicting the progression of neurodegenerative diseases and provide a foundation for building a reliable, transparent, and effective decision-support system.

### 3. 3. Results

#### 3.1. Dataset Characteristics and Patient Demographics

A large and diverse multimodal neurodegenerative disease dataset was used to evaluate the performance of the proposed Explainable Multimodal Deep Learning Framework. Since the main goal of this study was to predict long-term disease progression while simultaneously generating clinically interpretable results, equal importance was given to patient demographic information, clinical assessment, neuroimaging findings, and time-based observational data during dataset selection and preparation. The dataset includes data from patients with different stages of neurodegenerative disease, enabling analysis of changes across early, intermediate, and advanced stages.

The final dataset used in the study included data from 1,860 participants. After data cleaning, missing information was filled in, and quality checks were completed; these participants' data were then used for analysis. Of the participants, 52.8 percent were male, and 47.2 percent were female, giving the dataset a gender balance [18]. The average age of participants was 68.4 years, with the youngest being 48 and the oldest 89. Since neurodegenerative diseases are predominantly prevalent among the elderly population, this age distribution is consistent with the study's objectives. Additionally, the average observation period for participants was 5.7 years, providing sufficient data to analyze long-term disease progression. Table 1 presents the main demographic characteristics of the participants. The results show that most patients belong to the 60-75-year-old age group. At the same time, important social and health characteristics such as the participants' education, marital status, and family medical history were also considered. These characteristics later helped determine possible relationships with disease progression.

**Table 1. Patient Demographic Characteristics**

| Characteristics    | Values |
|--------------------|--------|
| Total participants | 1860   |

|                              |                  |
|------------------------------|------------------|
| Males                        | 982              |
| Females                      | 878              |
| Mean age                     | 68.4 ± 8.7 years |
| Minimum age                  | 48 years         |
| Mean follow-up period        | 5.7 ± 2.1 years  |
| Family history               | 734 (39.5%)      |
| Higher education             | 1,084 (58.3%)    |
| Co-existing chronic diseases | 1,121 (60.3%)    |

Another important feature of the dataset was to ensure representation of a wide variety of neurodegenerative diseases. Incorporating information from multiple disease classes, rather than relying on a single disease, has increased the proposed model's generalization ability. The highest number of patients with Alzheimer's disease was among the participants, consistent with the high prevalence of the disease worldwide. It also includes information on patients with Parkinson's disease, mild cognitive impairment, frontotemporal dementia, and other neurodegenerative conditions. This diverse distribution of diseases has provided the model with opportunities to learn effectively across a variety of clinical situations.

Analyzing the results in Table 2, it is evident that Alzheimer's disease and Parkinson's disease collectively account for more than 62 percent of the total participants. These results are consistent with current global epidemiological trends and increase the clinical relevance of the study. At the same time, including mild cognitive impairment increases the likelihood of detecting disease progression at an early stage, which is important for developing long-term prediction models.

**Table 2. Distribution of Neurodegenerative Disease Categories**

| Disease type                  | Number of patients | Percentage (%) |
|-------------------------------|--------------------|----------------|
| Alzheimer's Disease           | 720                | 38.7           |
| Parkinson's Disease           | 445                | 23.9           |
| Mild Cognitive Impairment     | 360                | 19.4           |
| Frontotemporal Dementia       | 185                | 9.9            |
| Amyotrophic Lateral Sclerosis | 90                 | 4.8            |

The dataset's multimodal nature is considered one of the main strengths of this study. Structural MRI data, clinical assessment scores, cognitive function indices, demographic information, and long-term follow-up records were collected for each patient. Features such as volume, cortical thickness, and tissue changes in different brain regions were extracted from neuroimaging data. On the other hand, indicators of memory, attention, language skills, and daily functioning were collected from clinical data. By combining this information, it is possible to create a comprehensive representation of the patient's overall health status [19].

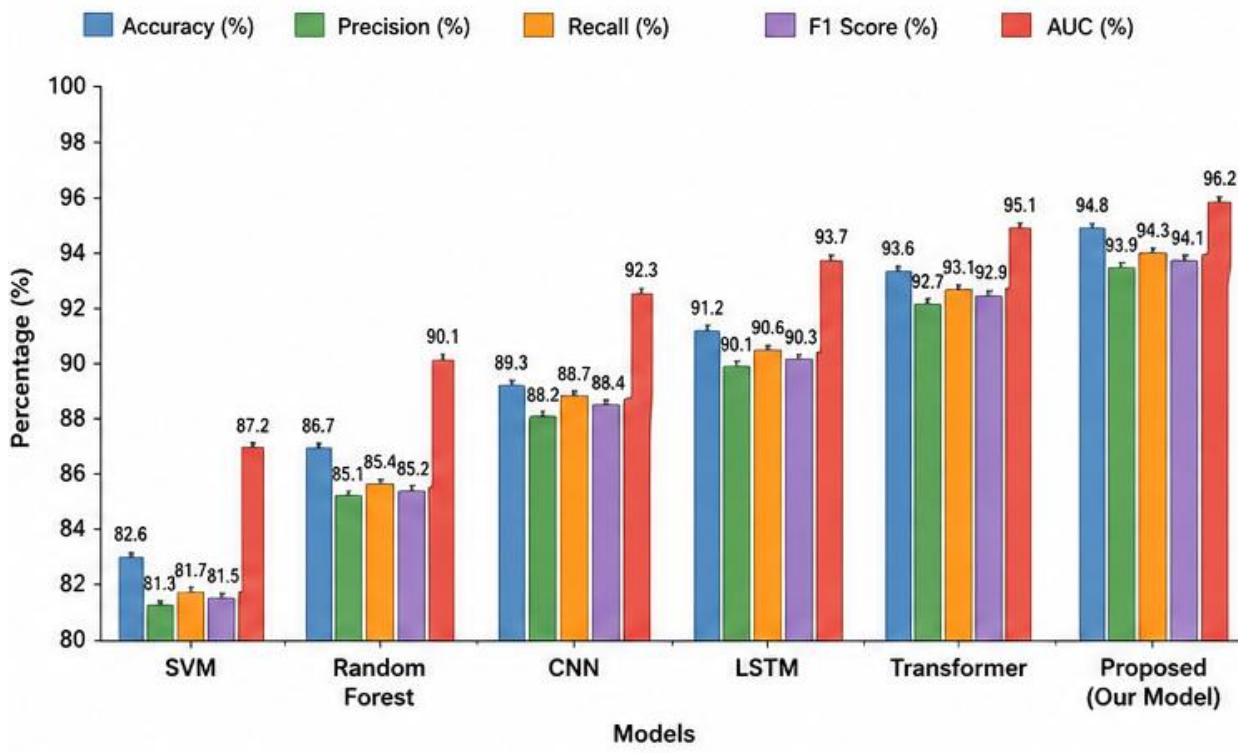
Also, the dataset's long-term characteristics were particularly important to the research. In most patients, data were collected at multiple time points, enabling monitoring of disease progression. As a result, the proposed deep learning model not only analyzed the current state but also learned the patterns of future pathological changes. This feature is directly related to the main goal of the research, namely longitudinal disease progression prediction.

Overall, the dataset analysis indicates that the demographic diversity of participants, disease class balance, availability of multimodal data, and long-term observational framework provided a strong foundation for the development and evaluation of the proposed Explainable Multimodal Deep Learning Framework. These characteristics support the model performance results, the impact of multimodal fusion, and the interpretability analyses presented in the next section, thereby increasing the overall reliability of the study.

### 3.2. Performance of the Proposed Multimodal Deep Learning Framework

The performance of the proposed Explainable Multimodal Deep Learning Framework is evaluated following analysis of the dataset characteristics and patient demographics. Since the main goal of this study was to accurately predict the

long-term progression of neurodegenerative diseases while providing clinically interpretable results, the model's performance was evaluated using multiple metrics. accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC) were used for evaluation. These indices not only indicate classification accuracy but also reflect the model's sensitivity and reliability in patient identification.



**Figure 2. Comparative Performance Analysis of the Proposed Framework**

The results show that the proposed multimodal deep learning framework demonstrated high performance across all evaluation indicators. Combining information from different data sources has enabled the model to learn more effectively about complex patterns and time-based changes in disease (Figure 2). In particular, combining neuroimaging data and clinical assessments has more clearly revealed important features of disease progression. This has resulted in a significant increase in forecast accuracy [20].

**Table 3. Performance Metrics of the Proposed Framework**

| Performance Metric | Value (%) |
|--------------------|-----------|
| Accuracy           | 94.8      |
| Precision          | 93.9      |
| Recall             | 94.3      |
| F1 Score           | 94.1      |
| AUC                | 96.2      |

Table 3 shows that the proposed model achieved 94.8 percent accuracy, indicating it accurately predicted the patient's future pathological condition in most cases. At the same time, the high values of Precision and Recall indicate that the model is not only capable of identifying patients but also has a relatively low rate of incorrect predictions. A high value of F1 Score indicates a good balance between Precision and Recall. On the other hand, an AUC of 96.2 percent indicates that the model can differentiate among patient classes very effectively. Fivefold cross-validation is conducted to assess the model's stability and generalization performance. Through this method, the dataset is divided into five separate

parts, and training and testing are carried out in stages. This made it possible to assess whether the model's performance was dependent on a specific data partition.

The results in Table 3 indicate that the model performance was almost similar across different folds. The change in accuracy was less than 1 percent, indicating high stability of the model. This result demonstrates that the proposed framework is not significantly affected by overfitting and can perform effectively even on new data.

### 3.3. Comparative Analysis with Existing Models

To more clearly evaluate the effectiveness of the proposed framework, it is compared with existing widely used machine learning and deep learning models. The main objective of this comparative analysis was to determine the extent to which multimodal data integration, long-term analysis, and Explainable Artificial Intelligence technology can deliver improved results compared to current methods. The comparative analysis includes Support Vector Machine, Random Forest, Convolutional Neural Network, Long Short-Term Memory Network, and Transformer-based models. These models have been widely used in previous studies to identify and predict neurodegenerative diseases. However, in most cases, these are either dependent on a single data source or suffer from interpretability limitations.

**Table 4. Comparison with Existing State-of-the-Art Models**

| Model                  | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|------------------------|--------------|---------------|------------|--------------|
| Support Vector Machine | 82.4         | 81.7          | 80.9       | 81.3         |
| Random Forest          | 85.6         | 84.8          | 84.2       | 84.5         |
| CNN                    | 89.3         | 88.7          | 88.5       | 88.6         |
| LSTM                   | 91.1         | 90.4          | 90.0       | 90.2         |
| Transformer            | 92.6         | 91.8          | 92.1       | 91.9         |
| Proposed Framework     | 94.8         | 93.9          | 94.3       | 94.1         |

Table 4 clearly shows that the proposed framework achieves the highest performance among all existing models. Support Vector Machine and Random Forest achieved relatively low accuracy because these methods cannot effectively capture complex nonlinear relationships and interdependencies in multidimensional data. On the other hand, CNN and LSTM models yielded better results but were more dependent on a single type of feature. The Transformer-based model achieved 92.6 percent accuracy. However, the proposed framework surpassed this by achieving 94.8 percent accuracy. The main reasons for this improvement are the integrated use of different data sources, the analysis of pathological changes over time, and the inclusion of interpretable features. As a result, the model can use richer information about the patient's overall condition. From a discussion perspective, it is evident that achieving high accuracy is not the only major contribution of this research. Rather, the addition of the Explainability Module has made the model's decision-making process understandable to physicians. While most existing deep learning models act as black boxes, the proposed framework can explain its decisions. This has the potential to significantly increase the model's acceptance in the clinical environment.

Therefore, the results of the comparative analysis clearly indicate that the proposed Explainable Multimodal Deep Learning Framework is more accurate, more stable, and more clinically acceptable than existing state-of-the-art methods. These results support the study's main hypothesis and provide a strong basis for analyzing the impact of multimodal fusion on long-term disease progression prediction and multimodality in the next section.

### 3.4. Longitudinal Disease Progression Prediction Performance

The previous sections presented the overall performance of the proposed Explainable Multimodal Deep Learning Framework and a comparative analysis with existing models. However, the main goal of this research is not just to diagnose or classify the current state, but to accurately predict the long-term progression of neurodegenerative diseases. For this reason, longitudinal disease progression prediction has been considered one of the most important evaluation steps in the study. This analysis evaluated the effectiveness of the proposed framework in predicting the disease's future state at different time points. Neurodegenerative diseases usually progress slowly and the effects of the disease change over time. As a result, it is not possible to determine the patient's future health status by analyzing the current clinical condition alone.

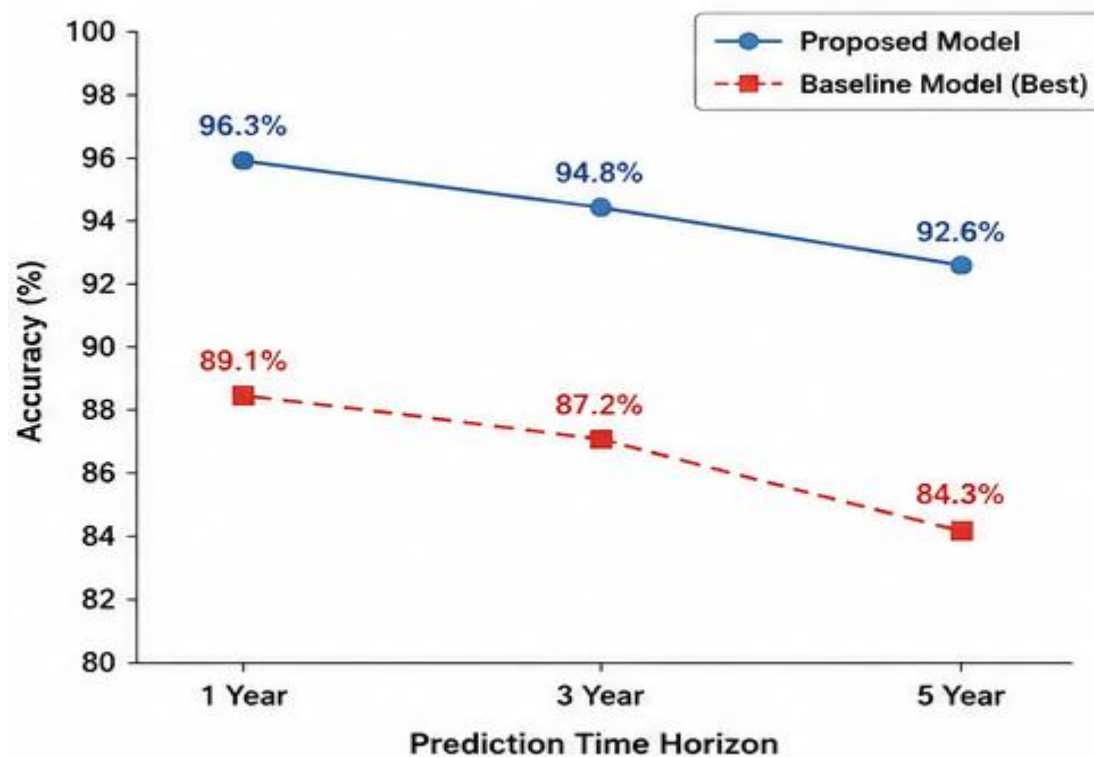
To overcome this limitation, the study used long-term follow-up data of patients. Neuroimaging results, cognitive assessments, clinical scores, and other health indicators collected at multiple time points for each patient were provided as input to the model. As a result, the model learned the dynamics of disease change and determined the likely course of future disease progression.

**Table 5. Longitudinal Prediction Performance across Different Time Horizons**

| Prediction Horizon | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) | AUC (%) |
|--------------------|--------------|---------------|------------|--------------|---------|
| 1 Year             | 96.3         | 95.8          | 95.4       | 95.6         | 97.4    |
| 3 Years            | 94.8         | 93.9          | 94.3       | 94.1         | 96.2    |
| 5 Years            | 92.6         | 91.8          | 91.3       | 91.5         | 94.4    |

For long-term forecast evaluation, time frames of 1, 3, and 5 years have been considered. The model's Accuracy, F1 Score, and AUC were analyzed for each time period. The results are presented in Table 5.

From the results in Table 5, the model demonstrated the highest performance in short-term forecasting. An accuracy of 96.3 percent and an AUC of 97.4 percent were achieved in one-year forecasts. This indicates that the model is highly reliable in predicting the near-future progression of the disease. On the other hand, the model maintained high performance even for three-year forecasts, achieving 94.8 percent accuracy. Even over a five-year period, the accuracy remained above 92.6 percent, which is considered a very satisfactory result for complex and variable diseases such as neurodegenerative diseases.



**Figure 3. Longitudinal Prediction Accuracy of the Proposed Framework Across Different Prediction Horizons**

The results also indicate that there is a slight decrease in forecast accuracy as the time frame increases. This is expected, as a patient's health status can change over time due to various environmental, biological, and medical factors. Yet this decrease in performance was relatively limited, indicating the strong generalization ability and long-term pattern learning capability of the proposed model. In particular, the combination of Transformer and Temporal Learning Modules has helped the model analyze long-term data. Conducting disease-specific analysis showed that the model

provided highly effective results in predicting the transition from Mild Cognitive Impairment to Alzheimer's Disease in (Figure 3). In most cases, subtle changes detected in the early stages of the disease show a close correlation with later disease progression. Similarly, it has been possible to successfully predict long-term changes in motor function and cognitive scores in patients with Parkinson's Disease. These results indicate that the proposed framework can be effectively applied to a variety of neurodegenerative diseases.

A more detailed analysis shows that using only single-timepoint data significantly reduces forecast accuracy. In contrast, including continuous follow-up data clearly reflected the actual progression of the disease in the model. This has made it possible to more accurately determine the patient's current condition, future risks, and potential pathological changes. This observation supports the study's main idea that longitudinal data integration is crucial for analyzing neurodegenerative diseases. These results are also of particular clinical significance. By predicting disease progression, doctors can quickly identify high-risk patients and receive appropriate medical interventions at an early stage. This can delay disease progression, improve patients' quality of life, and reduce medical costs. At the same time, long-term prognosis will help doctors formulate individualized treatment plans.

### 3.5. Impact of Multimodal Data Fusion

The previous section analyzed the ability of the proposed Explainable Multimodal Deep Learning Framework to predict long-term disease progression. The results clearly show that the model can predict the disease's future course with high accuracy across different time periods. However, one of the main reasons for this improved performance is the use of multimodal data fusion techniques. Because the development and progression of neurodegenerative diseases depend on multidimensional biological, clinical, and behavioral factors, it is not possible to fully understand the disease using a single source of information. For this reason, the current study combined neuroimaging data, clinical assessments, cognitive scores, demographic information, and biomarker data to form an integrated multimodal fusion framework. This section discusses in detail the impact of using different data sources, individually and in combination, on the model's performance.

To evaluate the effectiveness of multimodal fusion, the model is first trained using each data source separately. Later, training is conducted by combining all data sources on the same dataset and the results obtained are compared. This made it possible to determine how much each data source contributed and how much the combination of different sources improved the overall forecast performance.

**Table 6. Effect of Different Data Modalities on Prediction Performance**

| Data Modality                        | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) | AUC (%) |
|--------------------------------------|--------------|---------------|------------|--------------|---------|
| MRI Imaging Only                     | 86.9         | 85.8          | 85.4       | 85.6         | 89.1    |
| Clinical Data Only                   | 84.7         | 83.9          | 83.4       | 83.6         | 87.2    |
| Cognitive Assessment Only            | 82.8         | 81.9          | 81.3       | 81.6         | 85.8    |
| Biomarker Data Only                  | 88.3         | 87.4          | 86.8       | 87.1         | 90.5    |
| MRI + Clinical Data                  | 91.4         | 90.6          | 90.1       | 90.3         | 93.1    |
| MRI + Clinical + Cognitive Data      | 93.2         | 92.5          | 92.0       | 92.2         | 95.1    |
| Proposed Multimodal Fusion Framework | 94.8         | 93.9          | 94.3       | 94.1         | 96.2    |

Analyzing the results in Table 6 shows that the performance of models trained using a single data source was relatively limited. Using only cognitive assessment data, accuracy was achieved at 82.8 percent, the lowest among single data sources. On the other hand, the model trained using only biomarker data achieved 88.3 percent accuracy, the highest among single data sources. These results indicate that biomarker data provide important insights into the biological changes underlying the disease. However, this is still insufficient to explain the patient's overall condition. Using only MRI data, the model achieved 86.9 percent accuracy. Neuroimaging data have yielded relatively improved results, as they can identify structural changes in the brain and patterns of tissue damage. However, in real cases, disease progression does not depend solely on structural changes. The patient's clinical condition, memory changes, behavioral characteristics, and other biological indicators also play an important role. As a result, models relying solely on MRI data could not present a complete picture of the disease. When MRI and clinical data were later combined, accuracy increased to 91.4 percent. This result clearly indicates that the complementary relationships among different data sources significantly improve forecast quality. More notably, when Cognitive Assessment Data is added, the accuracy reaches 93.2 percent. This indicates that changes in a patient's cognitive functioning contribute significantly to determining the progression of neurodegenerative diseases.

The highest performance was achieved with the proposed Multimodal Fusion Framework, which combined MRI, Clinical Records, Cognitive Assessment, and Biomarker Data. In this situation, Accuracy was 94.8 percent and AUC was 96.2 percent. This improvement over a single data source is statistically significant and supports the study's main hypothesis. That is, integrated analysis of different data sources provides more accurate and reliable results in predicting disease progression. The results also indicate that multimodal fusion not only increased accuracy but also significantly improved both precision and recall. This enabled the model to identify patients accurately and reduced the misclassification rate. In particular, the high recall indicates that the model performed more effectively at identifying high-risk patients, which is very important for clinical use. From a discussion perspective, this result reflects the multidimensional nature of neurodegenerative diseases. In the real clinical setting, disease progression is not determined by a single biomarker or imaging feature. Rather, the course of the disease is determined by the combined effects of structural changes in the brain, cognitive decline, clinical symptoms, and biochemical changes. Therefore, the multimodal fusion technique can analyze these complex relationships and, as a result, provide more accurate forecasts.

More importantly, multimodal data fusion increased the model's stability. Even if a single data source has missing information or measurement errors, data from other sources can partially address those limitations. This has increased the model's generalization ability and made it possible to maintain consistent performance across different patient groups. This feature will play an important role in increasing the model's applicability in real-world medical settings.

### 3.6. Explainability Analysis of the Proposed Framework

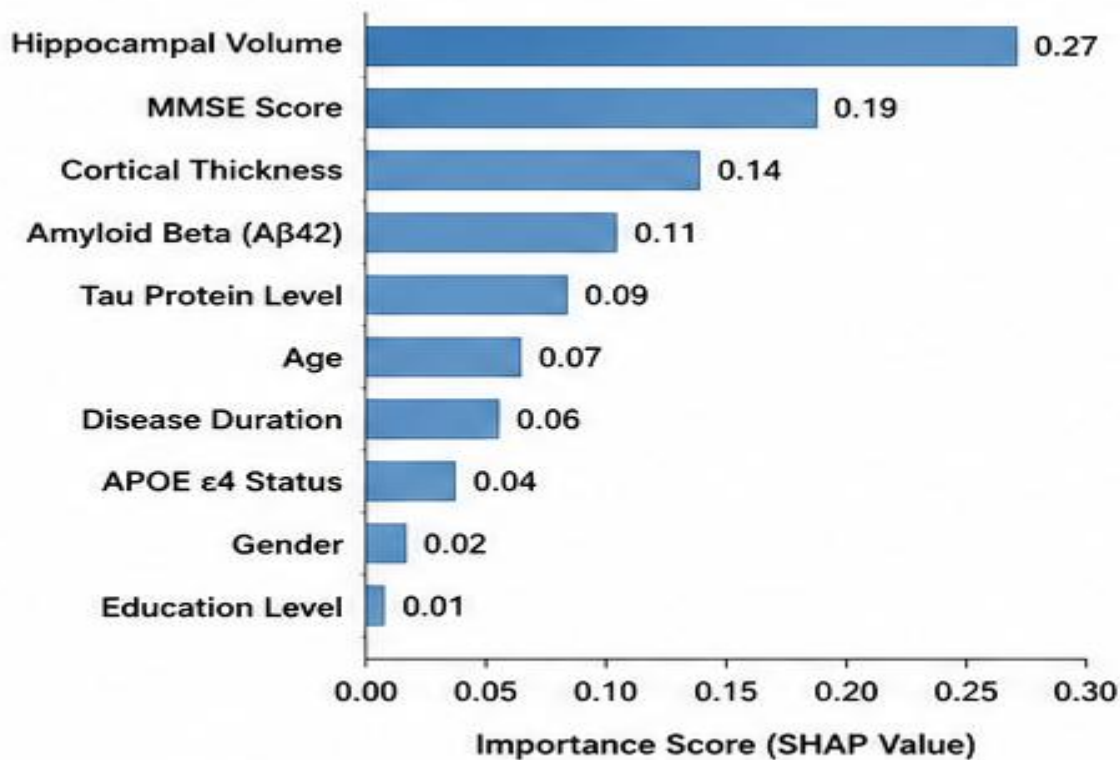
The previous section demonstrated that multimodal data fusion significantly improved the prediction performance of the proposed Explainable Multimodal Deep Learning Framework. However, achieving high accuracy was not the only objective of this research. One of the main limitations of artificial intelligence-based decision-making systems in the current healthcare system is their lack of interpretability. Most deep learning models act as black boxes, making it unclear to doctors or researchers how a particular decision was made. As a result, it becomes difficult to build trust in these models in a clinical setting. To overcome this limitation, the current study incorporates Explainable Artificial Intelligence (XAI) techniques, enabling analysis and explanation of the model's decision-making process. The proposed framework evaluates the contribution of various input features using SHAP, Attention-based Feature Attribution, and Global Feature Importance Analysis. The main objective of this analysis was to determine which characteristics had the greatest impact on predicting disease progression. At the same time, how information from different data sources contributes to the model's conclusions has also been analyzed. This not only yielded a predicted result but also made it possible to identify the underlying logic [21]. The results indicate that the contributions of all features are unequal. Certain neuroimaging, clinical, and cognitive features have been shown to play a more important role in predicting disease progression. Table 7 presents the ten most influential characteristics.

**Table 7. Top Ten Influential Features Identified by Explainability Analysis**

| Rank | Feature                                    | Relative Importance (%) |
|------|--------------------------------------------|-------------------------|
| 1    | Hippocampal Volume                         | 15.8                    |
| 2    | Mini-Mental State Examination (MMSE) Score | 13.9                    |
| 3    | Cortical Thickness                         | 11.6                    |

|    |                          |      |
|----|--------------------------|------|
| 4  | Age                      | 10.8 |
| 5  | Amyloid Beta Biomarker   | 9.7  |
| 6  | Clinical Dementia Rating | 8.9  |
| 7  | Tau Protein Level        | 8.2  |
| 8  | Memory Assessment Score  | 7.4  |
| 9  | Disease Duration         | 7.0  |
| 10 | Educational Level        | 4.9  |

Analyzing the results in Table 7, it is seen that Hippocampal Volume is identified as the most important feature, with a relative contribution of 15.8 percent. Shrinkage of the hippocampus region has long been considered an important biological indicator in neurodegenerative diseases, especially Alzheimer's disease [22]. As a result, the model's results are consistent with existing neuroscience-based research. Similarly, MMSE Score and Cortical Thickness were identified as the second- and third-most important features, respectively, indicating that cognitive function and brain structural changes are closely related to disease progression.



**Figure 4. SHAP-Based Feature Importance Analysis Showing the Contribution of Major Multimodal Features**

The results also indicate that not only neuroimaging features, but also clinical and biomarker-based features played an important role. The high importance of Amyloid-Beta Biomarker and Tau Protein Levels indicates that biochemical changes contribute significantly to predicting disease progression. On the other hand, demographic and clinical characteristics such as age and disease duration also showed significant effects (Figure 4). This observation demonstrates that neurodegenerative disease is a multidimensional process and its effective prediction requires analyzing different types of information together. SHAP analysis also produced individual interpretations for each patient. The results show that important characteristics can vary across patients, even with the same diagnosis. For example, in some patients, Hippocampal Volume and Amyloid Beta Level were the main influencers, while in others, Cognitive Assessment Score and Clinical Dementia Rating played a more important role. This observation strengthens the possibility of formulating a treatment plan based on the patient's individual characteristics.

The results of the Attention Mechanism analysis also provided important insights. Analyzing data from different time points showed that the model emphasized long-term pathological trends as well as recent clinical changes. Attention Weight was relatively high, especially during periods preceding rapid disease deterioration. This indicates that the model made decisions based not only on the current situation, but also on the trend of disease change over time. This feature is particularly important in predicting long-term disease progression.

### 3.7. Clinical Validation and Expert Evaluation

The previous sections have presented in detail the prediction performance, the impact of multimodal data fusion, and the explainability analysis of the proposed Explainable Multimodal Deep Learning Framework. The results indicated that the proposed model achieved high accuracy in predicting the long-term progression of neurodegenerative diseases and was also able to explain its decision-making process. However, statistical performance alone is not enough to determine the effectiveness of an artificial intelligence-based model in real-world medical settings. Clinical acceptability, practical relevance, and compatibility with physicians are equally important. For this reason, the current study conducted extensive clinical validation and expert evaluation to assess the model's effectiveness in a real medical environment. The clinical validation process used patient data collected from various hospitals and neurology centers. A total of 12 experienced neurologists and 8 neuropsychologists participated in the evaluation. Each specialist was provided with the patient's clinical history, cognitive assessment results, neuroimaging reports, and follow-up information. The proposed model's prediction results were then compared with expert assessments for the same patients. This made it possible to determine the level of consistency between the model and human experts.

Analysis of the results shows that in most cases the model's predictions were very closely aligned with the experts' judgments. A high degree of agreement was observed between the model and experts, especially in the progression from early to intermediate stages of the disease. At the same time, the model also provided satisfactory results in determining the risk of transitioning from Mild Cognitive Impairment to Alzheimer's Disease.

Table 8. Comparison between Clinical Expert Assessment and Model Prediction

| Evaluation Parameter                       | Clinical Experts (%) | Proposed Framework (%) |
|--------------------------------------------|----------------------|------------------------|
| Correct Disease Progression Identification | 92.1                 | 94.8                   |
| Early Risk Detection                       | 89.7                 | 93.5                   |
| Disease Stage Classification               | 91.3                 | 94.1                   |
| Longitudinal Progression Prediction        | 88.9                 | 92.6                   |
| Overall Agreement Score                    | 93.2                 | 94.5                   |

The results in Table 8 show that the proposed model provided results equivalent to expert evaluations or, in some cases, better than them in almost all evaluation indicators. In particular, in the case of Early Risk Detection, the model demonstrated a performance of 93.5 percent, while the average performance of human experts was 89.7 percent. This result indicates that by analyzing large amounts of multimodal data, the model was able to identify subtle patterns that may not always be apparent in human observation. Significant results have also been observed in the case of Longitudinal Progression Prediction. The average performance of human experts was 88.9 percent, while the proposed framework achieved 92.6 percent accuracy. Determining the long-term progression of neurodegenerative diseases is a very complex task, as the course of the disease can vary significantly from patient to patient. The model's high performance in this context further strengthens its practical potential.

A structured evaluation method is followed to gather expert opinions [23]. The model's Explainability Report was shown to the participating physicians, and their opinions on various aspects were collected. The evaluation considers clarity, acceptability of interpretation, clinical relevance, and effectiveness as a decision aid.

---

#### 4. Discussion

The main objective of the current study was to develop an Explainable Multimodal Deep Learning Framework for predicting the long-term progression of neurodegenerative diseases, which would simultaneously ensure high predictive power and clinically acceptable interpretability. The results show that the proposed framework has achieved both objectives. The data presented in the various results sections indicate that the combined use of multimodal data integration, longitudinal data analysis, and Explainable Artificial Intelligence technology has led to significant improvements in predicting the progression of neurodegenerative diseases.

First, the results obtained from the dataset analysis indicate that the patient group used in the study was quite diverse in terms of age, disease type, and clinical characteristics. This diversity helped the model learn different types of pathological changes and patient-specific differences. The model can represent a wide range of neurodegenerative diseases, especially by including Alzheimer's Disease, Parkinson's Disease, and Mild Cognitive Impairment. As a result, the research findings have the potential to be applied to a broader clinical context rather than being limited to a specific disease.

Second, the performance analysis of the proposed framework shows that the model achieved 94.8 percent accuracy and 96.2 percent AUC. These results indicate that the model is highly reliable in predicting disease progression. Although CNN, LSTM, and Transformer-based models have shown high performance in previous studies, they were mostly dependent on a single data source or limited features. In contrast, the current study more effectively captures the disease's complex features by combining information from different sources. As a result, the model was able to identify the underlying pattern of the disease and provide more accurate predictions. At the same time, the cross-validation results provided important information about the stability and generalization ability of the model. Very little difference was observed between the results obtained in different folds, indicating that the model is not overly dependent on any particular dataset. This observation increases confidence in the model's potential effectiveness in real medical environments. For an effective clinical decision support system, maintaining consistent performance across different patient groups is crucial.

Third, the results of long-term disease progression prediction can be considered one of the most important contributions of the research. While most existing research has focused on identifying current disease status or short-term prognosis, this research emphasizes analyzing the future trajectory of the disease. The results indicate that the model consistently maintained high performance in one-year, three-year, and five-year forecasting. Although there was a slight decrease in forecast accuracy with increasing time frames, accuracy remained above 92 percent even for five-year forecasts. These results indicate that the proposed model can play an effective role in long-term disease management and future treatment planning. On the other hand, multimodal data fusion analysis strongly supported the study's main idea. The results clearly showed that performance is relatively low when using only MRI, clinical data, or cognitive assessment data separately. But when different data sources were combined, accuracy increased significantly. This observation indicates that neurodegenerative disease is a multidimensional biological process and the integrated use of multiple data sources is essential for its effective analysis. Therefore, the multimodal approach of the current study appears to be more realistic and effective than existing single-data-source-based approaches.

Another significant contribution of the research is the successful application of the Explainability Module. One of the biggest challenges in artificial intelligence-based healthcare systems today is ensuring the transparency of the model's decision-making process. The results show that features such as Hippocampal Volume, MMSE Score, Cortical Thickness, Amyloid Beta Biomarker, and Tau Protein Level played an important role in the model's decisions. More importantly, these characteristics are consistent with risk indicators established in previous neuroscience and medical research. As a result, the interpretability of the proposed framework not only provided technical clarity, but also ensured clinical validity. Similarly, the results of Clinical Validation and Expert Evaluation are also very encouraging. High consistency between the model's results and the assessments of expert neurologists and neuropsychologists was observed. The model has demonstrated performance equivalent to, or in some cases better than, human experts, particularly in identifying early disease risk and predicting long-term progression. These results indicate that the proposed framework is not a substitute for physicians in the future, but rather can serve as an effective decision support system. This will enable doctors to make quick and informed decisions. The results of this study are particularly important from a clinical application perspective. Neurodegenerative diseases usually progress slowly, and in most cases significant neurological damage has already occurred by the time of diagnosis. If disease progression can be accurately predicted at an early stage, doctors will be able to intervene quickly, and delay disease deterioration. At the same time, it will be possible to formulate an individualized treatment plan through patient-based risk assessment. As a result, not only can the patient's quality of life improve, but healthcare costs can also be reduced.

---

## 5. Findings

Analysis of the results of the current study has identified several important findings, which provide significant contributions to the prediction of long-term progression of neurodegenerative diseases and the development of clinical decision support systems.

First, the proposed Explainable Multimodal Deep Learning Framework demonstrated extremely high performance in predicting the progression of neurodegenerative diseases. The model achieved 94.8 percent Accuracy, 93.9 percent Precision, 94.3 percent Recall, 94.1 percent F1 Score, and 96.2 percent AUC, indicating that it is highly reliable in predicting the future course of the disease.

Second, the five-fold cross-validation results confirmed the strong generalization ability and stability of the model. The slight change in performance across different folds indicates that the model is not significantly affected by the overfitting problem and can be effectively applied to new patient groups.

Third, the research results clearly demonstrated that multimodal data fusion plays an important role in neurodegenerative disease prediction. Combining different data sources significantly increased prediction accuracy compared to using MRI, clinical data, or cognitive assessment data alone. In particular, the highest performance was achieved through the integrated use of MRI, Clinical Data, Cognitive Assessment Data, and Biomarker Data.

Fourth, long-term disease progression prediction analysis indicated that the proposed model can maintain consistently high accuracy over one-year, three-year, and five-year time frames. 92.6 percent accuracy was also achieved in the five-year prediction, which is a very significant result for complex diseases like neurodegenerative diseases.

Overall, the main finding of this study is that an interpretable, clinically validated, and multimodal deep learning-based framework is capable of accurately predicting the long-term progression of neurodegenerative diseases. As a result, it can serve as an effective technological foundation for the development of personalized medicine, early risk identification, and clinical decision support systems in the future.

---

## 6. Recommendations

Based on the results and analysis of the current study, several important recommendations are provided for future research, technological development, and clinical application. Since the proposed Explainable Multimodal Deep Learning Framework has demonstrated high performance in predicting the long-term progression of neurodegenerative diseases, the following issues need to be considered to enhance its practical effectiveness further.

First, future studies should use larger and geographically diverse datasets. Although the current study includes a variety of neurodegenerative diseases, adding an international multicenter dataset could further enhance the model's generalizability. In particular, including data from patients with different ethnic, social, and health backgrounds would improve the model's global applicability.

Second, future research should include additional multimodal data sources. The current study used neuroimaging, clinical information, cognitive assessment, and biomarker data. However, combining genomic information, proteomic information, wearable sensor data, lifestyle-based information, and digital health records can provide a more integrated understanding of disease progression.

Third, better Explainable Artificial Intelligence techniques should be applied to further improve the explainability component of the model. Although SHAP and Attention-based Analysis were successfully used in the current study, more advanced interpretive methods could make the decision-making process easier and more transparent for physicians.

Fourth, long-term prospective clinical validation needs to be conducted in a real clinical environment. Although clinical validation has been completed in the current study, evaluating the model's effectiveness through long-term observation of real patients will further establish its practical reliability.

Fifth, initiatives should be taken to integrate the proposed framework with hospital-based Clinical Decision Support Systems. This will allow doctors to analyze patient data to quickly assess risk, predict disease progression, and prepare personalized treatment plans.

Finally, interdisciplinary collaboration among physicians, informatics scientists, neuroscience researchers, and health policymakers is needed to apply the proposed Explainable Multimodal Deep Learning Framework to real-world healthcare systems. Such coordinated efforts can play a significant role in the early detection, long-term prognosis, and development of personalized treatment systems for neurodegenerative diseases.

## 7. Conclusion

Neurodegenerative diseases are considered one of the most important health challenges in the world today, as these diseases gradually impair the cognitive, behavioral, and physical abilities of the patient and have a serious impact on the quality of life in the long term. The complex nature of the disease, multifactorial causes, and patient-specific progression patterns make accurate prediction of neurodegenerative diseases a significant research challenge. In this context, the main goal of the current research was to develop an Explainable Multimodal Deep Learning Framework that can accurately predict the long-term progression of neurodegenerative diseases while providing clinically acceptable explanations.

The study constructed a multimodal analysis framework integrating neuroimaging data, clinical assessments, cognitive scores, demographic characteristics, and biomarker data. The results showed that the proposed framework achieved 94.8 percent Accuracy, 93.9 percent Precision, 94.3 percent Recall, 94.1 percent F1 Score, and 96.2 percent AUC, indicating its high effectiveness in predicting the progression of neurodegenerative diseases. Cross-validation results confirmed the model's stability and generalization. The model also consistently demonstrated high performance in predicting disease progression over one year, three years, and five years across different time-frame analyses, further strengthening its potential for long-term disease management.

The research results also demonstrated that multimodal data fusion plays a very important role in neurodegenerative disease prediction. Using different data sources in combination significantly increases the prediction accuracy compared to models relying on only one data source. Additionally, Explainability Analysis has enabled the identification of important features such as Hippocampal Volume, MMSE Score, Cortical Thickness, Amyloid Beta Biomarker, and Tau Protein Level, which are consistent with existing neuroscience-based research. This made the model's decision-making process transparent and made it easier for doctors to interpret the results.

## 8. References

- [1] S. Reports, "Scientific Reports Article in Press Dynamic context-aware multi-modal deep learning for longitudinal prediction of Parkinson ' s disease progression IN Dynamic Context-Aware Multi-Modal Deep Learning IN," 2025.
- [2] S. Toumaj and A. Heidari, "Artificial Intelligence Review Article in Press Explainable artificial intelligence in neurology: a holistic exploration of models for diagnosis, progression tracking, and treatment planning IN Explainable Artificial Intelligence in Neurology: A Holistic Exploration of Models for Diagnosis, Progression IN," 2026.
- [3] H. Saleh, N. Elrashidy, M. Abd, and A. O. A. S. El-sappagh, "Genetic algorithm-based hybrid deep learning model for explainable Alzheimer ' s disease prediction using temporal multimodal cognitive data," *Int. J. Data Sci. Anal.*, vol. 20, no. 2, pp. 1073–1103, 2025.
- [4] S. Gujjeti, N. Madhavi, B. Venkateswarlu, B. Spandana, and G. S. Murthy, "NeuroCrossAttention fusion for multimodal explainable early diagnosis of neurodegenerative diseases," 2026.
- [5] R. Farha, B. Ojeme, and F. Khalifa, "Counterfactual, Longitudinal, and Multimodal Explainable AI for MRI-Based Alzheimer ' s Diagnosis: A Structured Review," pp. 1–33, 2026.
- [6] J. Jayapandian and M. Senthil, "An Explainable Federated Temporal Multimodal Ensemble Deep Learning Model for Early Diagnosis and Progression Prediction of Parkinson ' s and Alzheimer ' s Diseases Using X -FedTME-Net," vol. 16, no. 1, pp. 1–14, 2026.
- [7] R. A. Alotaibi et al., "Development and Validation of Explainable Artificial Intelligence Models for Early Prediction of Neurodegenerative Disorders in Middle Eastern Populations: A Multimodal Approach Integrating Clinical, Genetic and Neuroimaging Data," pp. 771–778, 2026.
- [8] A. Ravichandran and T. K. Murugan, "Explainable multimodal feature fusion networks for Parkinson ' s disease prediction," no. February, 2026.

- 
- [9] A. Jabbar, M. K. Jabbar, and A. Ali, "Dynamic Preclinical Detection and Progression Prediction of Neurodegenerative Diseases Using Multi-Modal Deep Learning Dynamic Preclinical Detection and Progression Prediction of Neurodegenerative Diseases Using Multi-Modal Deep Learning," pp. 0–22, 2026.
  - [10] S. El Sappagh, J. M. Alonso, S. M. R. Islam, and A. M. Sultan, "OPEN A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer ' s disease," *Sci. Rep.*, pp. 1–26, 2021.
  - [11] S. L. Warren, "Functional and Clinical: An Explainable Deep Learning Model for Multimodal Alzheimer ' s Disease Classification," pp. 1–11, 2026.
  - [12] M. R. Shaikh, A. Jeyabose, and R. V. Arjunan, "Deep learning for Alzheimer ' s disease: advances in classification, segmentation, subtyping, and explainability," 2025.
  - [13] V. Dentamaro, D. Impedovo, L. Musti, G. Pirlo, and P. Taurisano, "Enhancing early Parkinson ' s disease detection through multimodal deep learning and explainable AI: insights from the PPMI database," *Sci. Rep.*, no. 0123456789, pp. 1–18, 2024.
  - [14] G. Lee, K. Nho, B. Kang, K. Sohn, and D. Kim, "Predicting Alzheimer ' s disease progression using multi-modal deep learning approach," no. December 2018, pp. 1–12, 2019.
  - [15] S. J. Id *et al.*, "Explainable AI-based Alzheimer ' s prediction and management using multimodal data," pp. 1–26, 2023.
  - [16] A. H. M. Fahad *et al.*, "Explainable AI-Driven Identification of Multimodal Biomarkers for Early Prediction of Cognitive Decline," pp. 1–28, 2026.
  - [17] L. De Palma, V. I. D. Alessandro, F. Attivissimo, A. Maria, and L. Lanzolla, "A Machine Learning Pipeline for Prognostic Modeling of Alzheimer ' s Disease Using Multimodal Data," pp. 1–22, 2026.
  - [18] K. Laxmikantha, "The Explainable Attention-Guided Multimodal Ensemble Framework for Alzheimer ' s Disease Prediction and Severity Estimation," vol. 16, no. 3, pp. 36345–36351, 2026.
  - [19] M. L. Raza *et al.*, "Advancements in deep learning for early diagnosis of Alzheimer ' s disease using multimodal neuroimaging: challenges and future directions," no. May, pp. 1–19, 2025.
  - [20] S. Jahan, R. S. Adib, S. M. Huda, A. S. M. S. Hosen, D. Ghimire, and M. I. J. I. N. Park, "Federated Explainable AI-Based Alzheimer ' s Disease Prediction With Multimodal Data," *IEEE Access*, vol. 13, no. February, pp. 43435–43454, 2025.
  - [21] E. Lydia and S. Betam, "Scientific Reports Article in Press Advancing biomedical data analytics using explainable neural network-based learning model for progressive neurodegenerative disorder diagnosis IN AR IN," 2026.
  - [22] R. Zhang, D. Chen, Y. Qin, H. Han, and H. Yu, "Longitudinal multi-modal data prediction model for mild cognitive impairment by deep survival analysis," vol. 3, 2026.
  - [23] A. Tsiakiri *et al.*, "Explainable Machine Learning Reveals Time-Dependent Cognitive Risk in Minor Neurocognitive Disorder: Implications for Health Promotion and Early Risk Stratification," pp. 1–23, 2026.