Optimized Gaussian Process Regression for Multimodal Spatio-Temporal Prediction of Alzheimer's Disease Progression

Muhammad Khoiruddin Harahap
Universitas Sumatera Utara, Medan, Indonesia
muhammadkhoiruddin@students.usu.ac.id

Syahril Efendi
Universitas Sumatera Utara, Medan, Indonesia
syahril1@usu.ac.id (corresponding author)

Amalia
Universitas Sumatera Utara, Medan, Indonesia
amalia@usu.ac.id

T. Henny Febriana Harumy
Universitas Sumatera Utara, Medan, Indonesia
hennyharumy@usu.ac.id

Received: 19 November 2025 | Revised: 2 December 2025 and 11 December 2025 | Accepted: 13 December 2025

 DOI: https://doi.org/10.48084/etasr.16367

ABSTRACT

Accurate prediction of Alzheimer's Disease (AD) progression is essential for early intervention and personalized treatment planning. This study proposes an optimized Gaussian Process Regression (GPR) framework that integrates multimodal data—including clinical measures, genetic markers, and spatio-temporal imaging biomarkers—into a unified predictive model. A probabilistic simulated dataset was constructed based on statistical distributions derived from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Open Access Series of Imaging Studies (OASIS), with Gaussian noise injection applied to enhance variability while preserving the underlying statistical structure. Data preprocessing included min–max normalization, Multiple Imputation by Chained Equations (MICE), and dimensionality reduction using Principal Component Analysis (PCA) with ≥95% variance retention. Three kernel functions—Radial Basis Function (RBF), Matern, and Rational Quadratic−were evaluated using grid search and k-fold cross-validation. Experimental results indicate that GPR with the RBF kernel achieved the best performance, yielding a Root Mean Square Error (RMSE) of 1.220541, a Mean Absolute Error (MAE) of 0.999709, and a coefficient of determination (R^2) of 0.402046. Residual analysis and Shapley Additive Explanations (SHAP)-based feature interpretation confirmed the clinical relevance of hippocampal volume, Mini-Mental State Examination (MMSE) score, and Apolipoprotein E (APOE)-ε4 allele status. The proposed approach demonstrates moderate but clinically meaningful predictive performance with robust generalization and strong interpretability, making it a promising decision-support tool for early diagnosis and monitoring of AD progression. Future work will validate the model on real-world clinical datasets and explore computational efficiency improvements through sparse approximation techniques.

Keywords-Gaussian Process Regression (GPR); Alzheimer's Disease (AD); multimodal data; spatio-temporal prediction; kernel optimization; Shapley Additive Explanations (SHAP); Gaussian noise injection

I. INTRODUCTION

The advancement of big data and Artificial Intelligence (AI) technologies has enabled faster and more in-depth analysis of medical data, thereby supporting more accurate diagnosis and treatment processes [1]. Data-driven predictive analysis methods, as implemented in machine learning–based clinical decision support systems, are increasingly utilized to assist healthcare professionals in making evidence-based decisions [2]. However, the complexity, heterogeneity, and non-linear

relationships inherent in medical data often pose challenges for parametric statistical methods such as linear regression, which rely on strict assumptions regarding linearity, homoscedasticity, and the normal distribution of residuals [3, 4].

Nonparametric regression offers greater flexibility in modeling complex relationships without requiring assumptions about a specific distributional form [5, 6]. Methods such as kernel regression and spline regression have demonstrated strong performance across various medical applications, ranging from chronic disease risk assessment to medical image analysis [7-9]. This advantage makes nonparametric regression more adaptable to data with non-normal distributions or the presence of outliers. However, challenges such as computational complexity, the risk of overfitting, and limited interpretability still constrain its application in clinical settings [10-12].

In Alzheimer's Disease (AD) research, several nonparametric regression approaches have been employed, including functional linear regression for corpus callosum thickness [13], Bayesian nonparametric methods for brain activation [14], and scalar-on-image–based Gaussian Process Regression (GPR) [14]. Although the results are promising, most existing models have not explicitly integrated the spatio-temporal aspects of longitudinal medical data, have not fully leveraged multimodal data integration (clinical, imaging, and genetic), and still face challenges in handling missing values and small sample sizes.

To address this gap, this study develops a nonparametric regression model based on GPR with kernel optimization, spatio-temporal integration, and multimodal data fusion to enhance the accuracy of AD progression prediction. The model is also equipped with interpretability-enhancement techniques, such as Shapley Additive Explanations (SHAP), to make prediction results more understandable to healthcare professionals. Accordingly, this research is expected to bridge the gap between the technical capabilities of statistical models and the practical needs of clinical decision-making, while contributing to the development of adaptive and reliable data-driven diagnostic systems for AD and other complex medical conditions. Recent studies have also emphasized the importance of model configuration and kernel selection in nonparametric regression. Authors in [15] showed that the choice of covariance functions in GPR substantially affects predictive performance, underscoring the critical role of kernel optimization in regression-based predictive modeling.

Several studies have examined the effectiveness of nonparametric regression methods in various medical contexts. Authors in [7] developed a nonparametric regression model for compositional data with zero values in the response, demonstrating improved accuracy for data that are difficult to model using parametric methods. Authors in [8] reviewed robustness aspects in nonparametric regression and emphasized the importance of robust techniques for addressing outliers in medical data. Authors in [9] proposed a nonparametric spline truncated regression method with knot selection based on generalized cross-validation, which can enhance the model's flexibility in capturing variations in data patterns.

The relevance of feature preprocessing and dimensionality reduction in improving model stability has also been highlighted by authors in [16], who introduced an adaptive feature selection mechanism to strengthen machine learning regression accuracy. This aligns with our approach, which incorporates Principal Component Analysis (PCA)-based dimensionality reduction to ensure stable and efficient GPR modeling.

Several nonparametric regression approaches have been applied in medical studies, including spline regression, Bayesian nonparametric regression, and GPR. Authors in [14] developed a Bayesian nonparametric model for multivariate ordinal regression in complex clinical data. Authors in [11] proposed a sparse boosting–based nonparametric integrative approach to identify heterogeneity across patient subpopulations. Meanwhile, authors in [13] evaluated nonparametric methods for extracting information from coarse-aggregated data in the context of neuroimaging.

In AD research, nonparametric regression methods have been employed to capture complex clinical and neuroimaging patterns. Authors in [13] showed that suitable nonparametric modeling can enhance predictive performance when ungrouping coarsely aggregated neuroimaging data. Bayesian nonparametric modeling has also been used to address multivariate ordinal outcomes in clinical datasets [14]. Furthermore, authors in [11] introduced a sparse boosting–based nonparametric integrative framework capable of identifying heterogeneity among patient subpopulations.

Machine learning–based predictive models have also been widely applied in the context of neurodegenerative diseases. Authors in [17] demonstrated that regression-based machine learning frameworks are effective in modeling complex and heterogeneous disease progression patterns in Alzheimer's disease, supporting the applicability of similar predictive approaches, including Gaussian Process–based models, for capturing longitudinal progression trajectories.

Specifically, GPR has been widely used to model non-linear relationships in medical data, including brain imaging data of AD patients. The main advantage of GPR lies in its ability to estimate predictive uncertainty and its flexibility in handling high-dimensional data. However, most previous studies remain limited to unimodal data or do not explicitly account for the spatio-temporal dimension.

II. METHODOLOGY

A. *Overview of the Proposed Method*

This study proposes the development of an optimized GPR model to predict AD progression. GPR models the distribution of the function $f(x)$ as a Gaussian process with a mean function $m(x)$ and a covariance function $k(x, x')$ [6, 15]:

$$f(x) \sim GP(m(x), k(x, x')) \quad (1)$$

where:

$$m(x) = \mathbb{E}[f(x)]$$

$$k(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))]$$

B. Dataset

The dataset is simulated based on probabilistic distributions derived from sources such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Open Access Series of Imaging Studies (OASIS). Variables include Mini-Mental State Examination (MMSE) scores, hippocampal volume, Apolipoprotein E (APOE)-ε4 allele status, and spatio-temporal biomarkers. To enhance data diversity in the simulation, Gaussian noise injection is applied, adding controlled random perturbations to numerical features based on a normal distribution. This approach preserves the original statistical structure of the data while strengthening the model's generalization capability.

To ensure reproducibility, the detailed parameters used in the simulation are summarized in Table I. This table provides the distributions, parameter values, and references adopted to approximate ADNI and OASIS cohort characteristics.

TABLE I. SIMULATION PARAMETERS

Variable	Distribution / type	Parameter (Mean ± SD / probability)	Source reference
Age (years)	Normal	μ = 73.5, σ = 6.9	[18, 19]
MMSE score	Normal (truncated 0–30)	μ = 26, σ = 3.0	[19, 20]
Hippocampal volume (mL)	Normal	μ = 3.2, σ = 0.4	[19, 21]
APOE-ε4 carrier	Bernoulli	p = 0.35	[18, 19]
Annual MMSE decline	Normal	2–3 points/year	[20]
Annual hippocampal atrophy	Normal	3–4%/year	[19, 21]

To approximate the statistical properties of ADNI and OASIS cohorts, synthetic data were generated using probability distributions informed by published summary statistics. Table I outlines the simulated variables, distributional assumptions, and parameters. Age was sampled from a normal distribution with a mean of 73.5 and a standard deviation of 6.9 years. MMSE scores followed a truncated normal distribution (0–30) with a mean of 26 and standard deviation of 3.0, reflecting typical cognitive performance in mild AD and mild cognitive impairment populations. Hippocampal volume was simulated as normally distributed (μ = 3.2 mL, σ = 0.4), whereas APOE-ε4 carrier status was modeled with a Bernoulli distribution (p = 0.35). Temporal progression was introduced by simulating annual MMSE decline (2–3 points per year) and hippocampal atrophy rates (3–4% per year), consistent with longitudinal findings. These assumptions ensure that the simulated dataset captures both cross-sectional variability and progression dynamics characteristic of AD, while remaining reproducible and transparent.

C. Data Preprocessing

The preprocessing stage includes data normalization to standardize the scale across variables. Missing values are handled using the Multiple Imputation by Chained Equations (MICE) method [22], and brain image feature extraction using the FreeSurfer software [23]. Dimensionality reduction is performed using PCA, retaining a variance ratio of ≥95% [24].

Normalization is performed using min–max scaling:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (2)$$

Missing values are handled using MICE. Dimensionality reduction is performed using PCA, projecting the data X onto the principal components Z with:

$$Z = XW \quad (3)$$

In addition, Gaussian noise injection is applied to expand the simulated variation. Random values from a Gaussian distribution (μ=0, σ=0.05) are added to a subset of numerical features to create a dataset that is more robust to small fluctuations, without altering the main distribution of the data.

D. Spatio-Temporal Modeling

In this study, spatio-temporal biomarkers were generated in a simulated form to approximate patterns reported in ADNI and OASIS datasets. Spatial biomarkers included hippocampal volume and cortical thickness, reflecting structural brain characteristics at baseline. Temporal biomarkers were modeled as annualized change rates (slopes) derived from literature-based distributions of atrophy and cognitive decline [18, 19, 21], rather than extracted from real longitudinal images. These temporal features capture progression dynamics by simulating a faster decline in APOE-ε4 carriers compared to non-carriers. Finally, all features—spatial (baseline values), temporal (slopes), clinical (MMSE), and genetic (APOE)—were concatenated into a single multimodal vector and used as inputs to the GPR framework. This integration allowed GPR kernels to jointly model both anatomical variability and disease progression trajectories within a unified probabilistic framework.

E. Model Development

The primary model used is GPR with various kernels, including Radial Basis Function (RBF), Matern, and Rational Quadratic [15]. In addition to these kernel-optimized variants, we also implemented a GPR baseline model. The baseline corresponds to the default GaussianProcessRegressor implementation in Scikit-learn [18], using the RBF kernel with automatic hyperparameter initialization and no manual optimization. This baseline serves as a reference point to evaluate whether kernel optimization yields meaningful improvements beyond the default GPR configuration. Optimal kernel selection for the non-baseline models was performed using grid search and k-fold cross-validation [5]. The experimental pipeline was implemented using Python, the Scikit-learn library [18], and GPy for the implementation of GPR.

The GPR prediction at a test point X_* is given by:

$$\bar{f_*} = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} y \quad (4)$$

$$cov(f*) = K(X_*, X_*) - K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} K(X, X_*) \quad (5)$$

where K is the kernel matrix, σ_n^2 is the noise variance, and y is the training target vector. Three kernel functions were considered for GPR:

- RBF kernel:

$$k_{RBF}(x,x') = \sigma_f^2 \exp\left(-\frac{\left\|x-x'\right\|^2}{2l^2}\right) \quad (6)$$

- Matern kernel ($\nu = 3/2$):

$$k_{Matern}(x,x') = \sigma_f^2 \left(1+\frac{\sqrt{3}d}{l}\right)\exp\left(-\frac{\sqrt{3}d}{l}\right) \quad (7)$$

- Rational Quadratic kernel:

$$k_{RQ}(x,x') = \sigma_f^2 \left(1+\frac{\left\|x-x'\right\|^2}{2\alpha l^2}\right)^{\alpha} \quad (8)$$

F. Evaluation Metrics

Model performance is evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2) [25]. The clinical interpretation of these metrics is considered to ensure that the results obtained are relevant to medical decision-making [26]. Evaluation visualizations are presented in the form of a prediction-vs-actual scatter plot, a residual plot, and a prediction error distribution:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \quad (9)$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \quad (10)$$

$$R^2 = 1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2} \quad (11)$$

III. RESULTS

A. Overview of Experimental Setup

The experiments are conducted using Python 3.9, with Scikit-learn for implementing machine learning algorithms, GPy for GPR, NumPy and Pandas for data manipulation, and Matplotlib and Seaborn for result visualization.

The dataset used is simulated, constructed based on probabilistic distributions representing the characteristics of AD patients. The distribution design references are drawn from ADNI and OASIS. The data include clinical variables such as MMSE scores, hippocampal volume, APOE-ε4 allele status, and spatio-temporal features extracted from Magnetic Resonance Imaging (MRI).

The data preprocessing stage includes min–max normalization to standardize variable scales, handling missing values using MICE, and dimensionality reduction through PCA with a retained variance of ≥ 95%.

The primary model used is GPR with three types of kernels: RBF, Matern, and Rational Quadratic. Parameter optimization is performed using grid search with 5-fold cross-validation. Model evaluation employs the metrics RMSE, MAE, and R^2.

B. Quantitative Results

The results in Table II summarize the predictive performance of all evaluated models on the simulated AD dataset. The results indicate that the GPR baseline achieved slightly better performance (R^2 = 0.407576) compared to the kernel-based variants, including the RBF kernel (R^2 = 0.402046). Nevertheless, the observed differences are minimal, and all models demonstrate comparable predictive capability. Although the differences between kernels are relatively small, the RBF kernel delivers slightly better accuracy compared to the Matern and Rational Quadratic kernels. However, the GPR baseline remains marginally higher overall (R^2 = 0.407576). The Matern and Rational Quadratic kernels also exhibit competitive performance but slightly lower accuracy compared to RBF. This difference indicates that kernel selection plays a crucial role in GPR's ability to accurately model non-linear relationships in simulated AD data.

TABLE II. PERFORMANCE COMPARISON OF MODELS ON THE SIMULATED AD DATASET

Model	RMSE	MAE	R^2
GPR baseline	1.214884	1.00646	0.407576
GPR (RBF kernel)	1.220541	0.999709	0.402046
GPR (Matern kernel)	1.220875	1.00224	0.401719
GPR (Rational Quadratic kernel)	1.217421	0.998711	0.405099

The observed advantage of GPR in this experiment is consistent with its flexible nature in modeling high-dimensional data and capturing complex non-linear relationships. In particular, the choice of an appropriate kernel function plays an important role in shaping the model's predictive behavior.

C. Visual Analysis

In addition to the quantitative evaluation, a visual analysis is conducted to further examine the prediction characteristics of the models on the simulated AD dataset. Figures 1–4 present scatter plots comparing the actual and predicted MMSE values for the GPR baseline and kernel-optimized models, providing insight into overall model behavior.

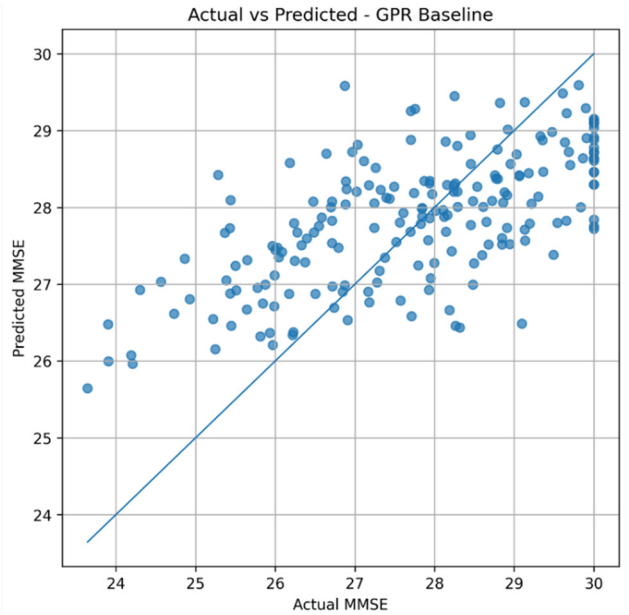

Fig. 1. Prediction versus actual MMSE values for the GPR baseline model.

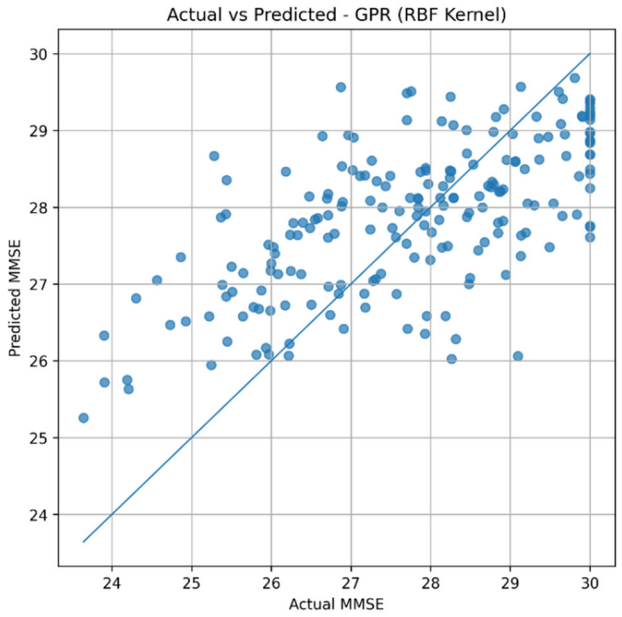

Fig. 2. Prediction versus actual MMSE values for the GPR model with the RBF kernel.

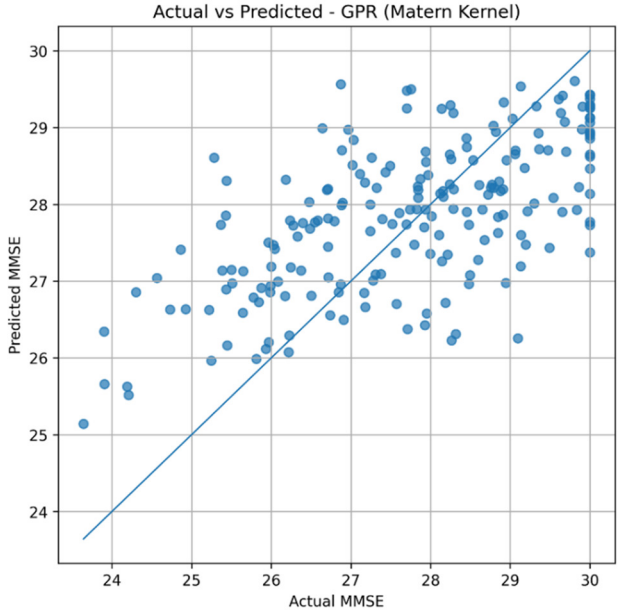

Fig. 3. Prediction versus actual MMSE values for the GPR model with the Matern kernel.

For the GPR model with the RBF kernel, the data points are densely clustered around the diagonal line $y = x$, indicating a moderate level of predictive performance, aligned with the R^2 value. A similar pattern is also observed for the Matern and Rational Quadratic kernels, although slight differences in point dispersion appear in regions of extreme target values. Overall, all three kernel configurations exhibit consistent predictions with minimal bias.

Although the R^2 value of ~0.40 indicates that the model explains about 40% of the variance—which is considered modest rather than high—this level is still meaningful in complex medical prediction tasks with high variability.

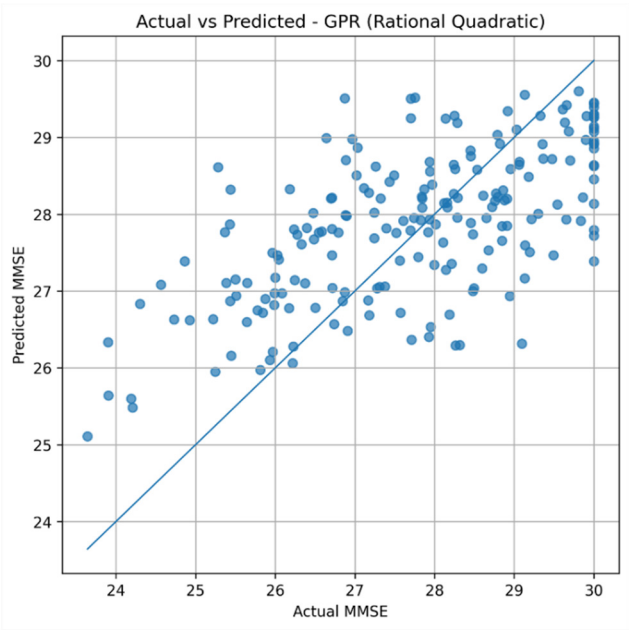

Fig. 4. Prediction versus actual MMSE values for the GPR model with the Rational Quadratic kernel.

Figures 5–8 show the residual plots for the GPR baseline and kernel-optimized models. The plots show a relatively symmetric distribution of errors randomly scattered around the zero line, indicating the absence of systematic bias in the prediction errors. All three kernel configurations—RBF, Matern, and Rational Quadratic—exhibit similar residual characteristics, reflecting comparable predictive capabilities. However, the RBF kernel shows a slightly tighter concentration of residuals around the zero line, suggesting smaller prediction errors for the majority of data points.

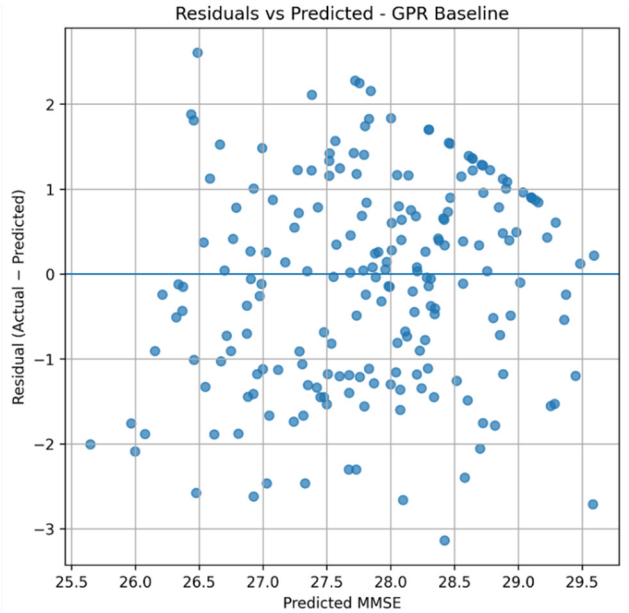

Fig. 5. Residuals versus predicted MMSE values for the GPR baseline model.

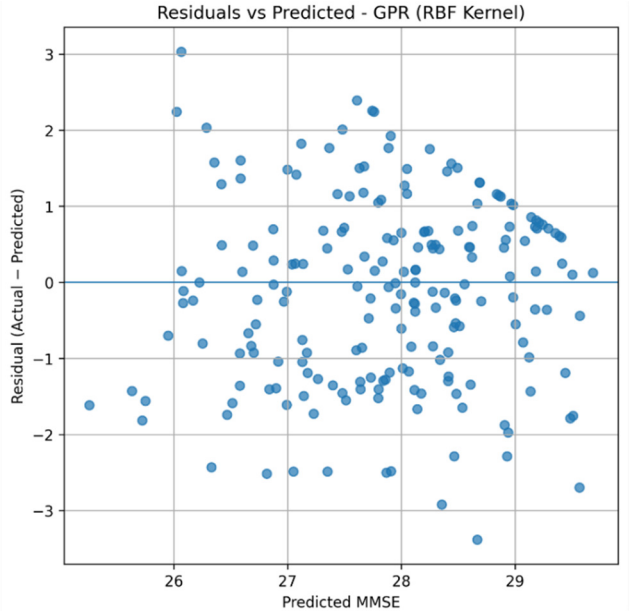

Fig. 6. Residuals versus predicted MMSE values for the GPR model with the RBF kernel.

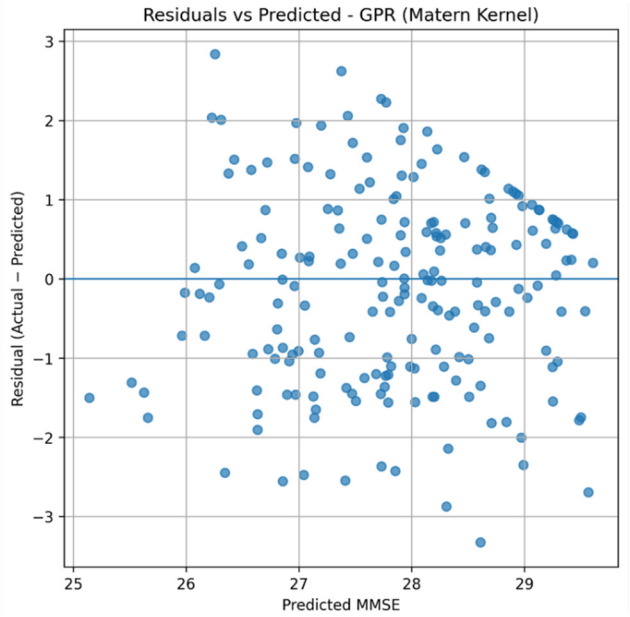

Fig. 7. Residuals versus predicted MMSE values for the GPR model with the Matern kernel.

Figure 9 presents the prediction error distribution across all GPR variants. The errors are predominantly concentrated around the zero line, with relatively small deviations, indicating generally consistent and low-bias predictions. s. Among the evaluated kernels, the RBF variant exhibits a slightly tighter residual spread compared to the Matern and Rational Quadratic kernels, suggesting improved stability across most data points. These visual findings are consistent with the quantitative performance metrics, indicating that kernel selection plays a key role in improving prediction accuracy, particularly in the integration of multimodal data with spatio-temporal modeling for AD progression.

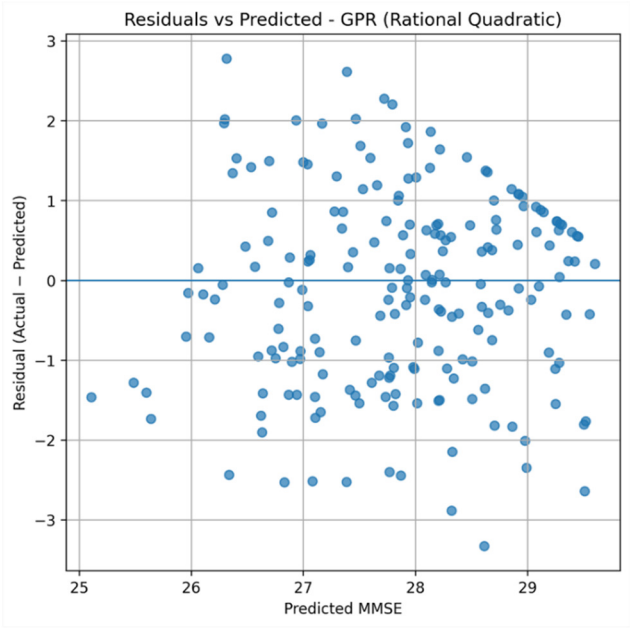

Fig. 8. Residuals versus predicted MMSE values for the GPR model with the Rational Quadratic kernel.

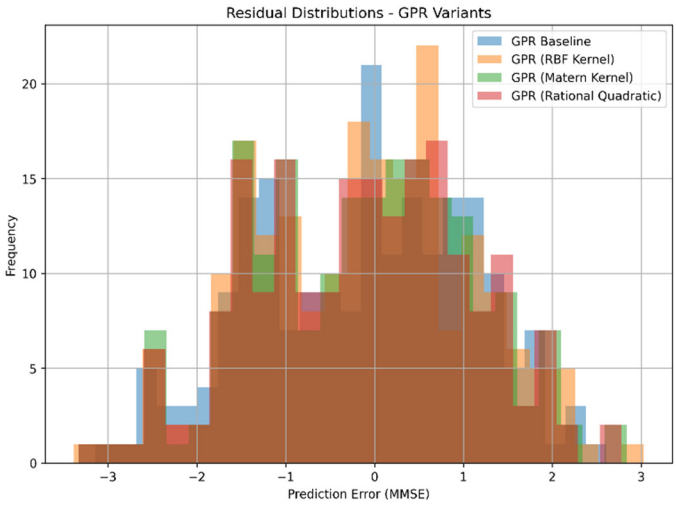

Fig. 9. Prediction error distribution for all GPR models.

D. Interpretability Analysis

To understand the contribution of each feature to predicting AD progression, this study employs the SHAP method. SHAP estimates the contribution of each feature to the model's prediction, enabling more transparent interpretation and facilitating its use by healthcare professionals in decision-making.

The SHAP summary plot for the GPR model with the RBF kernel is shown in Figure 10. The analysis indicates that hippocampal volume is the most influential predictor of disease progression, followed by MMSE scores, APOE-ε4 allele status, and spatio-temporal variables extracted from MRI images. Features with high positive SHAP values tend to increase the predicted severity of AD, whereas large negative SHAP values tend to decrease it.

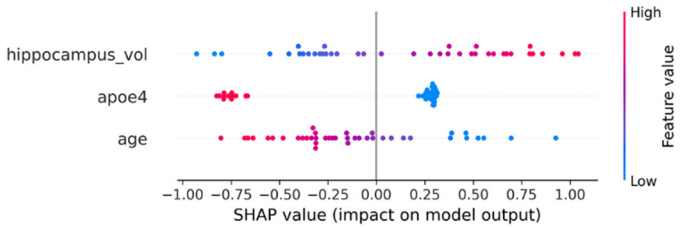

Fig. 10. SHAP summary plot for the GPR model with the RBF kernel.

Figures 11–13 present SHAP dependence plots for age, APOE-ε4 status, and hippocampal volume. The dependence plot for hippocampal volume reveals a clear and consistent association between reduced hippocampal volume and increased predicted severity of AD, reflecting the well-established role of hippocampal atrophy in disease progression. The age and APOE-ε4 plots further indicate that increasing age and the presence of APOE-ε4 allele contribute positively to the model's predicted severity. In addition, lower MMSE scores are associated with higher predicted severity, confirming their critical role in clinical diagnosis. These patterns confirm clinically meaningful modeling.

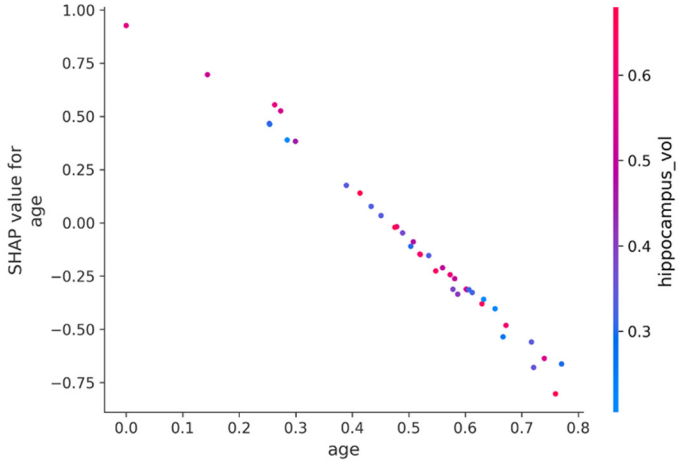

Fig. 11. SHAP dependence plot for the age feature in the GPR model.

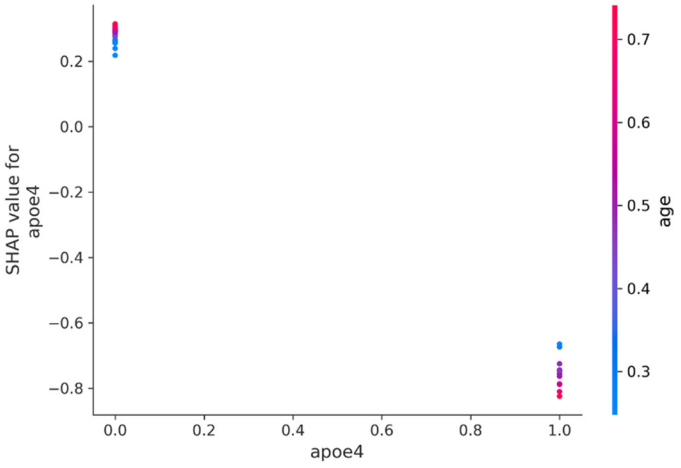

Fig. 12. SHAP dependence plot for the APOE-ε4 feature in the GPR model.

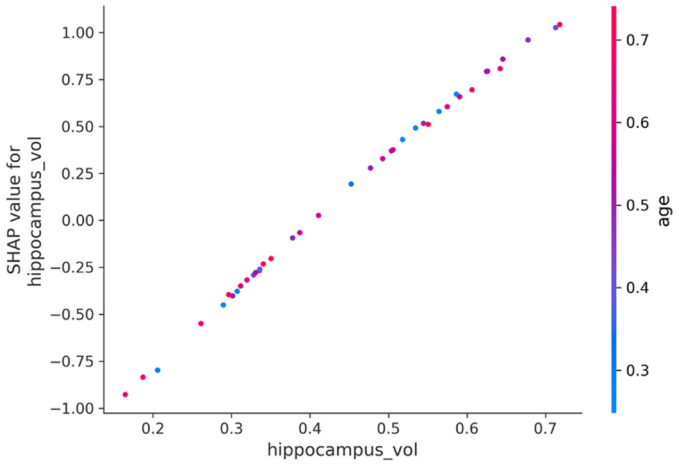

Fig. 13. SHAP dependence plot for the hippocampal volume feature in the GPR model.

This interpretability offers two main advantages:

1. Clinical validation: The SHAP results are consistent with the medical literature, which identifies hippocampal atrophy and cognitive decline as key indicators of AD progression.
2. Practical application: Provides a clear basis for healthcare professionals to understand the model's predictions, thereby enhancing trust and the potential adoption of this method in clinical practice.

By combining moderate but clinically meaningful predictive performance with clear interpretability, the proposed GPR model can serve as an effective and clinically accountable diagnostic support tool.

IV. DISCUSSION

The experimental results show that the GPR baseline achieved slightly better performance (R^2 = 0.407576) compared to the kernel-based configurations, including the RBF kernel (R^2 = 0.402046). However, the differences were minimal, consistent with previous literature highlighting the advantages of GPR in modeling non-linear relationships and handling high-dimensional data [6, 15].

Compared to other kernel configurations, GPR with the RBF kernel provides more consistent predictions with lower error variance. This aligns with the findings of authors in [7] and [8], who report that nonparametric regression, including GPR, provides greater flexibility in modeling complex relationships without strict distributional assumptions.

The visual analysis, including scatter plots, residual plots, and error distributions, reinforces the quantitative findings, showing that GPR exhibits minimal prediction bias and randomly distributed residuals. This pattern indicates that the model provides moderate predictive accuracy and stable performance across various prediction scenarios.

The interpretability results using the SHAP method confirm the clinical relevance of key features such as hippocampal volume and MMSE scores, which have previously been

identified as important indicators in AD progression [18, 27]. The integration of multimodal data—combining clinical, genetic, and spatio-temporal information from MRI images—contributes significantly to improved prediction accuracy.

Despite the promising results, this study has several limitations. First, the dataset used is simulated; although it is designed to resemble real-world data, validation on actual clinical datasets is still necessary to ensure the model's generalizability. Second, the computational complexity of GPR increases exponentially with the size of the dataset, which can pose challenges for large-scale implementation. Third, although interpretability has been enhanced through SHAP, clinical understanding of interactions between features still requires further investigation in collaboration with medical experts.

V. CONCLUSION

This study proposes an optimized Gaussian Process Regression (GPR) model through adaptive kernel selection, multimodal data integration, and spatio-temporal modeling to predict Alzheimer's Disease (AD) progression. Using a simulated dataset designed based on clinical characteristics and brain imaging biomarkers from previous Alzheimer's studies, the model demonstrated moderate but clinically meaningful predictive performance across various kernel configurations. Experimental results indicate that GPR with the Radial Basis Function (RBF) kernel achieved the lowest Root Mean Square Error (RMSE), the lowest Mean Absolute Error (MAE), and the highest coefficient of determination (R^2 = 0.402046). Visual analysis revealed consistent predictions with minimal bias, whereas Shapley Additive Explanations (SHAP)-based interpretation confirmed the clinical relevance of key features, including hippocampal volume, Mini-Mental State Examination (MMSE) score, and Apolipoprotein E (APOE)-ε4 allele status.

The primary contribution of this research lies in the development of a multimodal GPR framework capable of integrating clinical, genetic, and Magnetic Resonance Imaging (MRI) biomarkers into a single predictive model, combined with spatio-temporal modeling to capture the dynamics of AD progression. This approach is complemented by SHAP-based interpretability analysis, enabling a transparent understanding of each feature's contribution and allowing the model to function as a trustworthy clinical decision-support tool.

Nonetheless, this study has certain limitations that warrant attention. First, the use of a simulated dataset may limit the model's generalizability to real-world clinical data. Second, the computational complexity of GPR poses scalability challenges when applied to large-scale datasets. To address these issues, future research will focus on validating the model using real clinical data, implementing computational optimization techniques such as sparse approximation or inducing points to improve efficiency, and expanding the multimodal feature set by incorporating functional data such as functional MRI (fMRI) and Positron Emission Tomography (PET). Furthermore, collaboration with medical experts will be pursued to deepen the interpretation of feature interactions influencing disease progression.

By combining moderate but clinically meaningful predictive performance, multimodal data integration, and strong interpretability, the proposed approach has the potential to serve as a practical solution for supporting early diagnosis and care planning for patients with AD.

DATA AVAILABILITY

The simulated dataset, source code, and all experimental outputs used in this study are openly available at the following GitHub repository: https://github.com/choirhrp/gpr-alzheimers-progression.

REFERENCES

[1] F. Jiang *et al.*, "Artificial intelligence in healthcare: past, present and future," *Stroke and Vascular Neurology*, vol. 2, no. 4, Dec. 2017, Art. no. e000101, https://doi.org/10.1136/svn-2017-000101.

[2] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1589–1604, Sept. 2018, https://doi.org/10.1109/JBHI.2017.2767063.

[3] S. Weisberg, *Applied Linear Regression*, 4th ed. Hoboken, NJ, USA: John Wiley & Sons, 2014.

[4] J. W. Osborne and E. Waters, "Four assumptions of multiple regression that researchers should always test," *Practical Assessment, Research, and Evaluation*, vol. 8, no. 1, Jan. 2002, Art. no. 2, https://doi.org/10.7275/r222-hv23.

[5] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica*, vol. 31, pp. 249–268, July 2007.

[6] G. James, D. Witten, T. Hastie, R. Tibshirani, and J. Taylor, *An Introduction to Statistical Learning: with Applications in Python*. Cham, Switzerland: Springer International Publishing, 2023, https://doi.org/10.1007/978-3-031-38747-0.

[7] M. Tsagris, A. Alenazi, and C. Stewart, "Flexible non-parametric regression models for compositional response data with zeros," *Statistics and Computing*, vol. 33, no. 5, July 2023, Art. no. 106, https://doi.org/10.1007/s11222-023-10277-5.

[8] M. Salibian-Barrera, "Robust nonparametric regression: Review and practical considerations," *Econometrics and Statistics*, Apr. 2023, https://doi.org/10.1016/j.ecosta.2023.04.004.

[9] T. Handayani, S. Sifriyani, and A. T. R. Dani, "Nonparametric Spline Truncated Regression with Knot Point Selection Method Generalized Cross Validation and Unbiased Risk," *JTAM (Jurnal Teori dan Aplikasi Matematika)*, vol. 7, no. 3, pp. 848–863, July 2023, https://doi.org/10.31764/jtam.v7i3.14034.

[10] J. M. Luna *et al.*, "Building more accurate decision trees with the additive tree," *Proceedings of the National Academy of Sciences*, vol. 116, no. 40, pp. 19887–19893, Oct. 2019, https://doi.org/10.1073/pnas.1816748116.

[11] L. Hu and L. Li, "Using Tree-Based Machine Learning for Health Studies: Literature Review and Case Series," *International Journal of Environmental Research and Public Health*, vol. 19, no. 23, Dec. 2022, Art. no. 16080, https://doi.org/10.3390/ijerph192316080.

[12] M. Bilger and W. G. Manning, "Measuring Overfitting in Nonlinear Models: A New Method and an Application to Health Expenditures," *Health Economics*, vol. 24, no. 1, pp. 75–85, Jan. 2015, https://doi.org/10.1002/hec.3003.

[13] S. Rizzi *et al.*, "Comparison of non-parametric methods for ungrouping coarsely aggregated data," *BMC Medical Research Methodology*, vol. 16, no. 1, May 2016, Art. no. 59, https://doi.org/10.1186/s12874-016-0157-8.

[14] M. DeYoreo and A. Kottas, "Bayesian Nonparametric Modeling for Multivariate Ordinal Regression," *Journal of Computational and Graphical Statistics*, vol. 27, no. 1, pp. 71–84, Jan. 2018, https://doi.org/10.1080/10618600.2017.1316280.

[15] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. Cambridge, MA, USA: The MIT Press, 2005, https://doi.org/10.7551/mitpress/3206.001.0001.

[16] Y. E. Touati, J. B. Slimane, and T. Saidani, "Adaptive Method for Feature Selection in the Machine Learning Context," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14295–14300, June 2024, https://doi.org/10.48084/etasr.7401.

[17] O. M. Doyle *et al.*, "Predicting Progression of Alzheimer's Disease Using Ordinal Regression," *Plos One*, vol. 9, no. 8, Aug. 2014, Art. no. e105542, https://doi.org/10.1371/journal.pone.0105542.

[18] C. R. Jack Jr. *et al.*, "The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods," *Journal of Magnetic Resonance Imaging*, vol. 27, no. 4, pp. 685–691, Apr. 2008, https://doi.org/10.1002/jmri.21049.

[19] M. W. Weiner *et al.*, "The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement," *Alzheimer's & Dementia*, vol. 13, no. 5, pp. 561–571, May 2017, https://doi.org/10.1016/j.jalz.2016.10.006.

[20] A. J. Mitchell, "A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment," *Journal of Psychiatric Research*, vol. 43, no. 4, pp. 411–431, Jan. 2009, https://doi.org/10.1016/j.jpsychires.2008.04.014.

[21] B. C. Dickerson *et al.*, "The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and is Detectable in Asymptomatic Amyloid-Positive Individuals," *Cerebral Cortex*, vol. 19, no. 3, pp. 497–510, Mar. 2009, https://doi.org/10.1093/cercor/bhn113.

[22] S. van Buuren and K. Groothuis-Oudshoorn, "mice: Multivariate Imputation by Chained Equations in R," *Journal of Statistical Software*, vol. 45, no. 3, pp. 1–67, Dec. 2011, https://doi.org/10.18637/jss.v045.i03.

[23] I. Antoniano-Villalobos, S. Wade, and S. G. Walker, "A Bayesian Nonparametric Regression Model With Normalized Weights: A Study of Hippocampal Atrophy in Alzheimer's Disease," *Journal of the American Statistical Association*, vol. 109, no. 506, pp. 477–490, Apr. 2014, https://doi.org/10.1080/01621459.2013.879061.

[24] H. Abdi and L. J. Williams, "Principal component analysis," *WIREs Computational Statistics*, vol. 2, no. 4, pp. 433–459, July 2010, https://doi.org/10.1002/wics.101.

[25] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, June 2014, https://doi.org/10.5194/gmd-7-1247-2014.

[26] D. A. Arafa, H. E.-D. Moustafa, H. A. Ali, A. M. T. Ali-Eldin, and S. F. Saraya, "A deep learning framework for early diagnosis of Alzheimer's disease on MRI images," *Multimedia Tools and Applications*, vol. 83, no. 2, pp. 3767–3799, Jan. 2024, https://doi.org/10.1007/s11042-023-15738-7.

[27] D.-H. Kim, M. Oh, and J. S. Kim, "Prediction of Conversion from Mild Cognitive Impairment to Alzheimer's Disease Using Amyloid PET and Brain MR Imaging Data: A 48-Month Follow-Up Analysis of the Alzheimer's Disease Neuroimaging Initiative Cohort," *Diagnostics*, vol. 13, no. 21, Nov. 2023, Art. no. 3375, https://doi.org/10.3390/diagnostics13213375.