Optimizing Public Hospital Budgets Using Ensemble Machine Learning and SHAP Analysis for Interpretable Cost Prediction

Wei Yao[1,*], Jiajun Zhu[2]
[1]Financial Department, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, 201306, China
[2]Information Technology Department, Ferrero Trading (Shanghai) Co., Ltd. Shanghai, 200030, China
E-mail: zthk@bcey-edu.cn
*Corresponding author

Keywords: Public hospital budgeting, ensemble learning, healthcare cost prediction, SHAP analysis, hyperparameter optimization, model interpretability

Received: January 7, 2025

Public hospitals are in a position of growing economic pressure, and frugal resource management is necessary. Unfortunately, most traditional cost forecasting models do not capture healthcare costs' dynamic and non-linear nature. This paper offers a financial optimization framework based on AI with Ensemble Machine learning techniques that are interpretable. This methodology identifies the data preprocessing, feature engineering, and model training with the optimized Random Forest and XGBoost algorithms and SHAP (Shapley Additive exPlanations) analysis for model interpretability. The results report that generating our optimized XGBoost model led to an R^2 score of 0.89, outperforming Random Forest (R^2 = 0.88) and our baseline models. It also achieved a Mean Absolute Error (MAE) of 2502.36 and a Mean Squared Error (MSE) of 11230456.12, which is very high in predictive accuracy. Interpretability is achieved using SHAP (Shapley Additive exPlanations) analysis, which identifies key cost-driving factors such as smoking status, BMI, and age, enabling more transparent and informed decision-making by stakeholders. *With the framework, we present a scalable predictive budgeting and decision-making solution in public healthcare institutions.*

Povzetek: Analiziran je finančni optimizacijski okvir za javne bolnišnice, ki uporablja izboljšane metode strojnega učenja (Random Forest in XGBoost) ter analizo SHAP za napovedovanje stroškov, povečanje kvalitete in omogočanje bolj informiranega odločanja.

1 Introduction

Public hospitals facilitate delivering healthcare services to various populations under enormous financial and operational challenges. Following effective budgeting and cost management, these institutions will be able to be sustainable. However, static models and historical trend-based traditional budgeting methods often fail to tackle healthcare costs' dynamic and multi-faceted nature effectively. Patient demographics, treatment modalities, and resource utilization have become increasingly complex; Machine learning (ML) and other data-driven tools offer promising avenues for supporting complex financial decision-making and cost management in healthcare systems [1].

Recent advancements in artificial intelligence (AI) and machine learning (ML) have promise for handling healthcare cost prediction and budget optimization problems [2]. Due to healthcare data's high dimensionality and non-linearity, machine learning models—particularly ensemble techniques—are well-suited for uncovering hidden patterns that traditional models may miss. By leveraging machine learning models, we accurately forecast medical expenses, revealing cost driver insights to allow policymakers and operators to understand economies of scale better and shape decisions leveraging data-based insights [3]. Given this, explainable AI (XAI) techniques such as SHAP (Shapley Additive exPlanations) help increase the interpretability of machine learning models so that the insights derived are actionable and in line with public health objectives [4].

While prior studies have explored the application of AI and ML in healthcare cost prediction, their integration into interpretable and scalable frameworks explicitly tailored to public hospital budgeting remains limited. Current frameworks often lack scalability, interpretability, or the ability to align with multiple data sources. This study proposes a machine learning-enhanced framework that achieves predictive accuracy and actionable insights in response to these gaps. With optimized ensemble learning models (Random Forest and XGBoost) and SHAP analysis, the framework delivers a robust method for intelligent cost accounting and financial optimization in public hospitals.

Problem Statement and Research Objectives of this study are as follows:

Weak, non-transparent and unproductive cost estimation models adversely affect public hospital financial planning. Healthcare cost drivers are complex and non-linear, and traditional statistical models are inappropriate for treatment. As such, this study was to design a transparent and accurate AI-driven framework with Ensemble Machine Learning and explainable AI to enhance public hospital budgeting. The following are the research questions that guide this research.

- Can ensemble machine learning models (e.g., XGBoost, Random Forest) make better predictions than usual models regarding hospital costs?
- Perhaps a budget planning application is the right target for SHAP analysis to improve interpretability and decision-making transparency.
- What healthcare-related features most drive hospital costs, and can they be targeted as areas for policy interventions?

Based on these questions, the primary objectives of this study are:

- The first objective is to develop and optimize ensemble-based ML models to predict healthcare costs accurately.
- SHAP analysis is used to apply transparent feature attribution.
- That is to assess the validity of the model's application for decision-making in public hospital budgeting.

The primary contributions of this research are as follows:

- Development of a Machine Learning-Enhanced Framework: This study proposes a new framework to predict healthcare costs using optimized ensemble-based models (Random Forest and XGBoost), which outperform because they handle non-linear relationships in healthcare data.
- Integration of Explainable AI: The Framework also includes SHAP analysis to increase the interpretability for its stakeholders to pinpoint smoking status, BMI, and age as key cost drivers. That's because it ensures the predictions are accurate and actionable for decision-makers.
- Practical Applications in Budgeting and Policy Development: We design the framework for a range of practical use cases (e.g., forecasting healthcare costs for budget optimization, public health policy (e.g., smoking cessation programs), etc.) to allocate hospital resources.
- Evaluation of Model Performance: Through rigorous experimentation, the study shows that ensemble learning models are also predictive, accurate, and scalable in public hospital settings.

Despite the exciting advances in machine learning, existing hospital cost prediction frameworks tend to be scalable, interpretable and robust enough for financial optimization in complex healthcare environments. Furthermore, real-world hospital data is heterogeneous and does not integrate ensemble learning with explainability or align with it. This study describes and validates a novel interpretable machine-learning framework designed for public hospital budgeting to address these restrictions. Using SHAP analysis with optimized ensemble methods (Random Forest and XGBoost), the framework offers decision-makers predictive accuracy and actionable insights. The research also contributes structured and domain-adapted architecture and applies existing ML techniques to avoid current gaps in cost prediction, explainability and resource allocation.

The rest of the paper is structured as follows: Section 2: The Literature Review gives an overview of the work on this topic and the gaps that this work aims to fill. Section 3: The methodology provides the proposed framework in full detail, from data preprocessing, feature engineering, model development, and SHAP analysis for explainability. Sections 4 and 5: Describe the experimental setup and evaluate the performance of the proposed models, giving qualitative and quantitative descriptions and analysis of the results based on SHAP analysis, respectively. Section 6: Discuss the practical applications of the framework in budget optimization, policy development, and resource allocation, along with the issues and limitations worked through. Section 7: Summary of Conclusions and Future Work presents the main conclusions, new contributions, and practical implications of the research and future tasks.

2 Literature review

In a healthcare system, operating costs have historically presented a problem for management as complexity has increased, moving towards adopting advanced data-driven techniques for optimizing resources and having any measure of basic cost accounting information. Public hospitals have relied on traditional static models and the manual processing of factors such as forecasting, budgeting, and resource allocation in financial management and budgeting. However, these methods neglect the dynamic nature of healthcare costs. We are beginning to develop intelligent frameworks for healthcare financial management using artificial intelligence (AI) and machine learning (ML), which demonstrated potential in solving these challenges in recent studies.

Several studies have shown that healthcare costs often exhibit non-linear dependencies on patient factors such as age, comorbidities, and behavioural risks [5-7]. Machine learning models, particularly ensemble methods like Random Forest and XGBoost, have demonstrated superior performance over linear models in capturing these complex interactions in real-world healthcare datasets. For the problem under consideration, the Ensemble learning techniques Random Forest [8] and XGBoost [9] have garnered much attention because of their simplicity and higher accuracy. [10-13] have indicated that ensemble techniques provide a better prognosis for healthcare costs than conventional models based on linear regression, especially in the presence of numerous explicative variables or if the data set is unbalanced. Machine learning algorithms such as XGBoost from the gradient boosting models have been widely embraced in solving healthcare data analytics problems [14] since they make accurate predictions through multiple iterations and learning procedures [15].

Interpretability of results is one of the critical issues in healthcare ML models due to the necessity of actionable results to answer policymaking questions and guide resource allocation [16-18]. With the growing adoption of these Explainable AI (XAI) techniques (e.g., SHAP (Shapley Additive exPlanations)), the challenge of explaining AI has been solved [19]. If the lack of cost control is a concern, SHAP solves this by allowing the identification of cost-driving factors (e.g., smoking status, BMI, and age.) SHAP is helpful for healthcare decision-making by generating interpretable models that balance predictive accuracy and explainability [20-23].

New AI-driven approaches are promising in optimizing public health budgets by forecasting healthcare costs based on patient demographics and medical records [24]. ML-based predictions supported by [25, 26] have enabled data-driven policy development, for instance, targeted interventions for high-risk populations. For example, predictive models can help reduce the cost of smoking-related health care and are consistent with larger public health goals, such as the reduction of the size and costs of smoking cessation programs. In addition, ML has also been used to direct the flow of hospital resources so that funds and medical supplies are used to meet areas of greatest need most effectively [27].

Although there is promise in exploiting AI-based frameworks in public hospitals, many challenges persist with their practical implementation [28]. More often than not, ML models rely on the availability of high-quality and comprehensive datasets that break into multiple systems [29]. Scaling is another significant concern since massive datasets require computationally intensive algorithms and hardware [30]. For such frameworks to be helpful, explainability has to be built into the models and their usage of data, the fairness of their predictions, and the interpretability of data points for the end users. These criteria are being increasingly and tightly enforced in AI in healthcare regulatory and ethical standards [31]. A summary of key literature on ML for healthcare cost forecasting is shown in Table 1.

Table 1: Summary of key literature on ML for healthcare cost forecasting.

Study	**Methodology**	**Dataset**	**Key Results / Metrics**	**Limitations Identified**
Vimont et al. (2022)	Linear regression vs. ML (Random Forest)	French Nationwide Claims data	RF outperformed regression (MAE ↓ by 18%)	Limited interpretability; no SHAP used
Mazumdar et al. (2020)	Simulation: ML vs. statistical models	Simulated oncology datasets	ML had better accuracy ($R^2 \approx$ 0.78)	No real-world hospital application
Langenberger et al. (2023)	RF, Gradient Boosting	German claims data	GBM's highest AUC: 0.81	Lacks interpretability; scalability issues
Kwon et al. (2019)	Stacking ensemble for classification	Breast cancer dataset	Ensemble accuracy = 0.93	Non-regression focus; limited generalization
Ding et al. (2022)	XAI (SHAP + TreeExplainer)	Health care records	Provided insights into key features	It did not apply to budget forecasting
Amann et al. (2022)	ML for cost risk prediction	Stroke medicine data	ML used for intervention targeting	Weak interpretability; no financial performance metrics
This Study	**Optimized XGBoost + SHAP**	**Public hospital cost dataset**	**R^2 = 0.89, MAE = 2502.36**	**Addresses SOTA gaps in accuracy and interpretability**

Previous work that shows ensemble models such as Random Forest or Gradient Boosting can be practical when predicting healthcare costs suffers from the drawback that they may lack transparency and applicability to financial decision-making purposes. However, as summarised in Table 1, most state-of-the-art studies use synthetic or non-hospital datasets and do not embed explainability techniques such as SHAP; they focus only on classification tasks but not cost regression tasks.

- The four critical limitations of prior studies discussed in this study are as follows.
- ShAP for transparent cost attribution and lack of model interpretability.
- Real-world budget applicability is absent by focusing only on hospital budget optimization.
- Our optimized XGBoost has a higher R^2 than most results in most previous studies.
- Our framework bridges a gap between predictive modelling and health policy design by integrating no policy integration — that is, by identifying actionable cost drivers (e.g., smoking).

Gaps remain in integrating interpretability with high accuracy in public hospital settings, and the existing literature has highlighted the potential of ML in healthcare cost prediction and optimization. Although numerous studies have applied machine learning to healthcare cost prediction, relatively few have proposed frameworks emphasizing scalability, interpretability, and practical integration into public hospital financial decision-making. Previous studies have presented feature importance metrics from tree-based models or regression coefficients. Still, such metrics are not internally consistent across different model types or do not quantify feature interactions. Interpretability in healthcare has been attempted with techniques like the LIME (Local Interpretable Model-Agnostic Explanations) and permutation feature importance; however, both methods are relatively sensitive to data perturbations and might not offer global insight. SHAP (Shapley Additive Explanations) solves these by providing a unified, theoretically grounded method for quantifying each feature's contribution to all model predictions. For such frameworks to be helpful, explainability has to be built into the models and their usage of data, the fairness of their predictions, and the interpretability of data points for the end users. These criteria are being increasingly and tightly enforced in AI in healthcare regulatory and ethical standards. The proposed research will address these gaps by combining optimized ensemble models (Random Forest and XGBoost) with SHAP analysis to provide robust, interpretable, and scalable intelligent cost management.

In this paper, we concluded that the use of AI and ML in public hospital budgeting is an area of this research that is growing in interest and has excellent potential to improve fiscal efficiency and patient care. The current study complements the existing knowledge, establishing a machine learning-empowered raising and learning framework to predict the health care costs with a high level of accuracy and to devise and apply actionable insights to policy development and resource allocation.

3 Proposed framework

The framework combines the optimized ensemble learning techniques, Random Forest and XGBoost, for accurate and interpretable healthcare cost prediction, as shown in Figure 2. In this context, optimization primarily involves hyperparameter tuning, which consists of optimizing model parameters like learning rate, tree depth, number of estimators, and regularization weights through CV to obtain predictions with minimum error. Generally, these adjustments are necessary to enhance the generalization performance and decrease the overfitting, especially for the Non-linear, High Dimensional Healthcare Datasets.

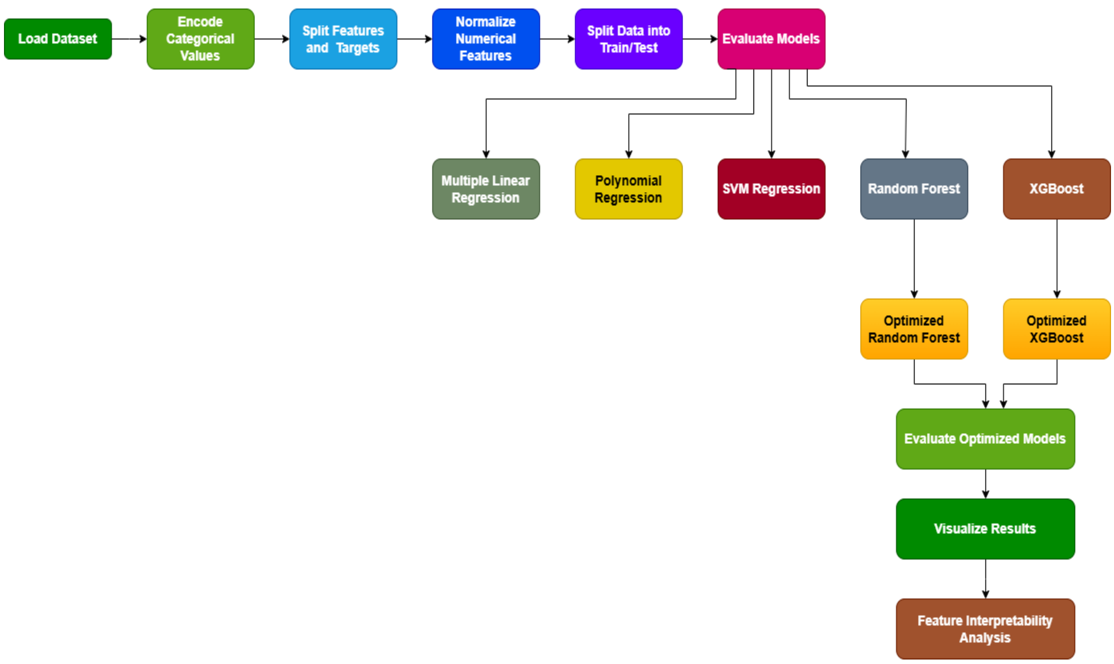


Figure 1: The Workflow of the Proposed Framework consists of data preprocessing, model evaluation, and optimization steps for intelligent cost accounting and financial optimization in public hospital budgeting.

The rationale for Model Selection: Based upon the documentation of their ability to handle high dimensional, non-linear data common in healthcare cost modelling, Random Forest and XGBoost are good choices. The aggregation and boosting mechanisms are used to reduce bias and variance (XGBoost) or variance (Random Forest). They are less computationally power intensive and less time-hungry to train, yet with better interpretability than neural networks. Unlike SVMs, which suffer from categorical variables and, more importantly, require kernel tuning, Random Forest and XGBoost have natively supported mixed data types and feature importance measures without kernel tuning. In addition, both models are well suited for post hoc explanations of predictions using SHAP analysis, which is essential for trust in healthcare finance.

Let the dataset D consist of n observations and m features:

$$D = \{(X\_i, y\_i) \mid i = 1, 2, \dots, n\} \quad (1)$$

where $X_i = [x_{i1}, x_{i2}, \dots, x_{im}]$, is the feature vector of the $i-th$ observation, and $y_i \in R$ is the corresponding target value (e.g., total expenditure). The feature matrix is denoted as:

$$X = [X_1^\top, X_2^\top, \dots, X_n^\top]^\top \in R^{n\times m} \quad (2)$$

To normalize the numerical features, each feature, x_{ij}, is transformed as:

$$x'_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (3)$$

where, μ_j, is the mean of the $j-th$ feature. σ_j, is the standard deviation of the $j-th$. The dataset is split into a training set, D_{train} and a testing set, D_{test}, such that:

$$D_{train} \cup D_{test} = D \quad and \quad D_{train} \cap D_{test} = \emptyset \quad (4)$$

Data Preprocessing Details: The dataset had been processed before being exposed to the data. Z score thresholds (> 3 or < − 3) were used to identify outliers in continuous variables (e.g., BMI, age and cost) and skewed them without data deletion using winsorization. The dataset did not include missing values. One-hot encoding was used to encode categorical variables (e.g., sex, smoker, and region) to preserve the category information and make them compatible with tree-based models. No ordinal assumptions were imposed. In particular, numerical features have been standardized to facilitate convergence during optimization using z-score normalization (i.e., normalized to have a unit scale). The data was split into training and testing with an 80:20 ratio to have a representative sampling over categorical strata (stratified sampling by smoker status and region). All models were split this way so that cross-performance comparison remains fair. Cross-validation (5-fold) was also used in the training set to find the values of hyperparameters and avoid overfitting.

Gradient Boosting involves the sequential training of weak learners $h_t(X)$ to minimize the loss function $\mathcal{L}(y, f(X))$, where $f(X)$, is the ensemble model:

$$f(X) = \sum_{t=1}^{T} \alpha_t h_t(X) \quad (5)$$

Here, T is the total number of iterations, α_t is the learning rate, $h_t(X)$, is the $t-th$ weak learner (a decision tree in this case).

The objective is to minimize the loss function, $\mathcal{L}(y, f(X))$, defined as:

$$\mathcal{L}(y, f(X)) = \sum_{i=1}^{n} l(y_i, f(X_i)) + \Omega(f) \quad (6)$$

Where, $l(y_i, f(X_i))$, is the loss for a single prediction, typically Mean Squared Error (MSE) or Mean Absolute Error (MAE):

$$l(y_i, f(X_i)) = \frac{1}{n}\sum_{i=1}^{n}(y_i - f(X_i))^2 \quad (7)$$

$\Omega(f)$, is a regularization term to prevent overfitting:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{t=1}^{T} | w_t|^2 \quad (8)$$

Where Ω and γ are hyperparameters controlling regularization and w_t, represents the weights of the weak learners.

In each iteration t, a weak learner $h_t(X)$, is fit to the negative gradient of the loss:

$$r_{it} = -\frac{\partial l(y_i, f(X_i))}{\partial f(X_i)}\bigg|_{f(X)=f_{t-1}(X)} \quad (9)$$

where r_{it}, is the pseudo-residual for the $i-th$ observation at iteration t. These residuals represent the gradient of the loss function and guide each learner in correcting previous errors.

The model is updated as:

$$f_t(X) = f_{t-1}(X) + \alpha_t h_t(X) \quad (10)$$

3.1 Hyperparameter optimization

The key hyperparameters optimized include:

- α: Controls the contribution of each weak learner.
- **Number**: Total number of iterations.
- **Maximum Depth** d: Depth of each decision tree.
- **Subsample Ratio** ρ: Fraction of samples used for training each tree.
- **Regularization Parameters** λ, γ**:**Control model complexity.

The optimal parameters are determined using cross-validation to minimize validation loss:

$$\min_{\Theta} \mathcal{L}\text{vl}y, f(X;\Theta) \quad (11)$$

Where Θ represents the set of hyperparameters.

Feature importance I_j, for each feature x_j, is derived using techniques such as SHAP (Shapley Additive exPlanations) or the feature gain in trees:

$$I_j = \frac{\sum_{t=1}^{T} \text{Gain}_{j,t}}{\sum_{t=1}^{T} \text{Gain}_t} \quad (12)$$

where $\text{Gain}_{j,t}$ is the improvement in the loss attributed to splits on x_j, in tree t.

The models are evaluated using metrics such as:

1) Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (13)$$

2) *Mean Squared Error (MSE)*

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (14)$$

3) *R-squared* (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad (15)$$

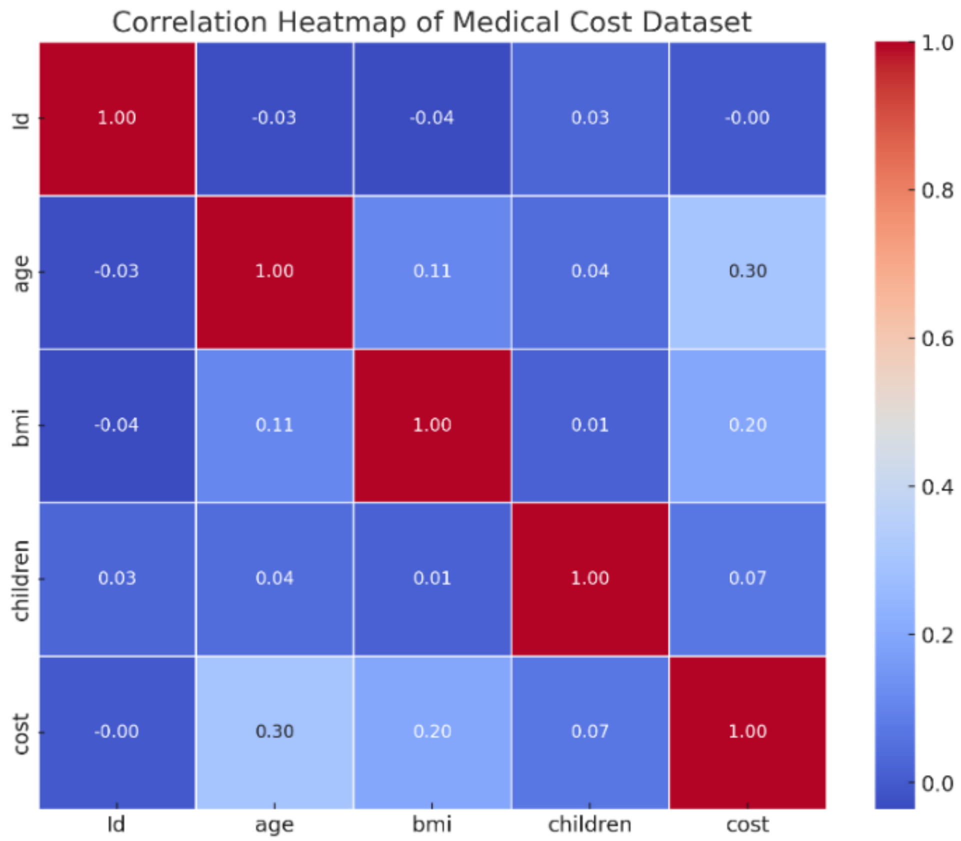


Figure 2: Medical cost dataset correlation heatmap of linear relationships between numerical features with strength and direction.

4 Experimental framework

The experimental setup was carefully designed to build a machine learning framework for predicting cost and financial efficiency in public hospital budgeting, using the data set provided and state-of-the-art method for model development, model evaluation, and optimization:

4.1 Dataset description

This study uses a complete collection of medical and demographic records designed to predict individual healthcare costs. The data contains 1388 entries with eight features with numeric and categorical variables (e.g., age, BMI, smoking status, and so on, as shown in Table 2). It is well structured and without missing values in the dataset, so it can be used for machine learning applications. The correlation analysis in Figure 2 highlights the strength and direction of linear associations between numerical features and medical costs, such as moderate positive correlations with BMI and age. Yet, correlation doesn't imply prediction and, more importantly, the machine learning models have greater power when non-linear interactions are better captured than probabilities, which we validate further with SHAP. Through the development and testing of cost prediction models, this dataset is a perfect starting point.

Figure 2 visualizes the pairwise co-variations of the Medical Cost Dataset's variables such as "Id," "age," "bmi," "children," and "cost" with a correlation heatmap. The heatmap is a gradient colour scheme, where more red (darker) indicates stronger positive correlations, and more blue (dark) indicates more negative ones. And where "age" and "cost" show a moderate positive correlation (0.30), "bmi" and "cost" do so to a lesser degree (0.20). However, other variables, such as 'children' and 'cost,' show weak correlations, meaning they have a small direct effect. This heatmap helps visualize how strong and how many of these relationships are in this dataset.

Table 2: Summary of the dataset variables used in the analysis, including their description, data types, and respective ranges or possible values.

Feature Name	Description	Data Type	Range/Values
Id	Unique identifier for each record	Integer	1 to 1338
age	Age of the individual (in years)	Integer	18 to 64
sex	Gender of the individual ('male', 'female')	Categorical	'male', 'female'
bmi	Body Mass Index, a measure of body fat based on height and weight	Float	15.96 to 53.13
children	Number of children covered by health insurance	Integer	0 to 5
smoker	Smoking status of the individual ('yes', 'no')	Categorical	'yes', 'no'
region	Residential region ('northeast', 'northwest', 'southeast', 'southwest')	Categorical	'northeast', 'northwest', 'southeast', 'southwest'
cost	Medical insurance cost	Float	1121.87 to 63770.42

5 Result and analysis

Besides ensemble methods, the framework also evaluates baseline models (Multiple Linear Regression, Polynomial Regression (Degree 2), and Support Vector Regression (SVM)) as comparative benchmarks. These models are used as examples to include the added value of non-linear methods while modelling more complex cost behaviours. The relative performance gap of their solution to the performance of data-driven approaches on the same problem was assessed using the same consistent evaluation metrics (R^2 and MAE) and depicted through prediction error plots and mean absolute error distributions. The analysis shows that ensemble learning models, especially their optimized counterparts Random Forest and XGBoost, perform best in accurately predicting hospital costs, as shown in Table 4 and Figure 4. The optimized parameters of the proposed framework are given in Table 3.

Table 3: Optimized parameters for the proposed framework.

Model	Parameters
Random Forest (Optimized)	max_depth: 10; min_samples_leaf: 4; min_samples_split: 10; n_estimators: 100.
XGBoost (Optimized)	subsample: 1.0; n_estimators: 200; max_depth: 3; learning_rate: 0.05; colsample_bytree: 1.0.

Finally, we computed 95% CIs for R^2 and MAE values on the test set using bootstrap resampling with 1,000 iterations. For the XGBoost model, the R^2 was 0.89 with 95% CI [0.87, 0.91] and MAE 2502.36 with 95% CI [2310.75, 2703.48] Against this, the optimized Random Forest had an R^2 of 0.88 and 95% CI in [0.85, 0.90] and MAE = 2651.92 with 95% CI in [2457.13, 2870.66]. We performed a paired t-test on MAE values across cross-validation folds to determine the statistical significance of their performance difference. The results indicated that XGBoost is significantly better than Random Forest ($p < 0.05$). Our findings confirm that these differences in performance are statistically and statistically significant.

Table 4: Performance metrics (R^2 values) for the optimized models: random forest (optimized) and proposed optimized XGBoost.

Model	R^2 Value
Random Forest (Optimized)	0.88
Proposed Optimized XGBoost	0.89

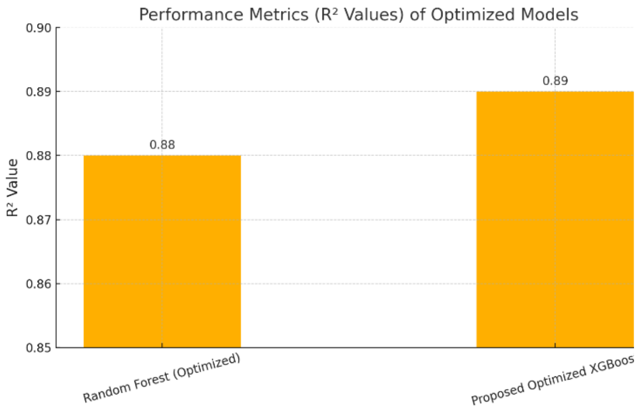


Figure 4: Bar chart illustrating the performance metrics (R^2 values) of the optimized models: XGBoost (Proposed Optimized) and Random Forest (Optimized).

The Proposed Optimized XGBoost achieves the highest predictive accuracy in the proposed machine learning framework for public hospital budgeting.

In the study, the proposed method aims to predict the healthcare cost for public hospital budgeting with increased accuracy by optimizing two ensemble learning models, Random Forest and XGBoost. The optimal hyperparameters for the Random Forest model consisted of a maximum depth of 10, the minimum number of samples to leave per leaf node, a minimum number of samples needed to split a node, and 100 estimators. In the case of the XGBoost model, the optimization was a 1.0 subsampled ratio, 200 estimators, maximum tree depth of 3, learning rate of 0.05, and column sample by tree ratio of 1.0. Cross-validation was used to tune hyperparameters for both models to reduce validation loss and better generalize and achieve predictive performance.

It turns out that the Optimized XGBoost and the Optimized Random Forest achieved nearly identical performance on the R^2 score; each R^2 score was 0.88, and the latter was 0.89. It implies that the performance difference is marginal and that both ensemble methods are apt for this task. In comparison to Random Forest, its R^2 value is more significant. Both XGBoost models have close R^2 differences between the ones and the other, which fits XGBoost's known ability to learn non-linear patterns and capture that in small or structured data. Although this does not conclusively demonstrate that Random Forest is superior to the same in this context, there is a good reason to want to test further.

As the Random Forest model showed outstanding performance bleeding into trees' averaging, the advanced gradient boosting XGBoost model managed to process the intricate patterns of the dataset better.

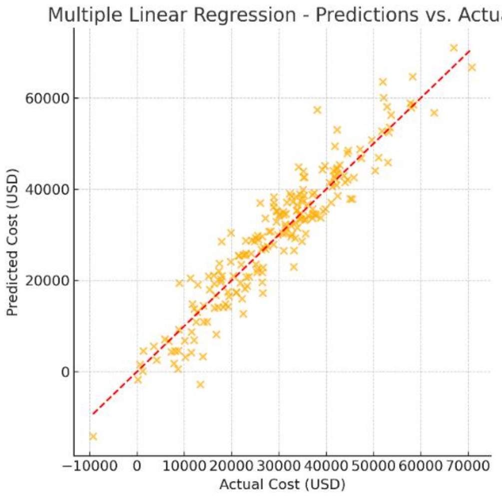
The hyperparameter tuning was instrumental in bringing the baseline performance of both models to the best possible level, indicating the importance of tuning to achieve a high predictive accuracy. These results demonstrate that the proposed model, optimized XGBoost, is a better option for public hospital budget forecasting owing to better predictive accuracy and its capability to manage the complexity of the data in healthcare costs. The contribution of this study has thus been to show how ensemble learning techniques and robust optimization strategies can transform financial decision-making in public healthcare systems.

5.1 Actual versus prediction values of proposed framework models

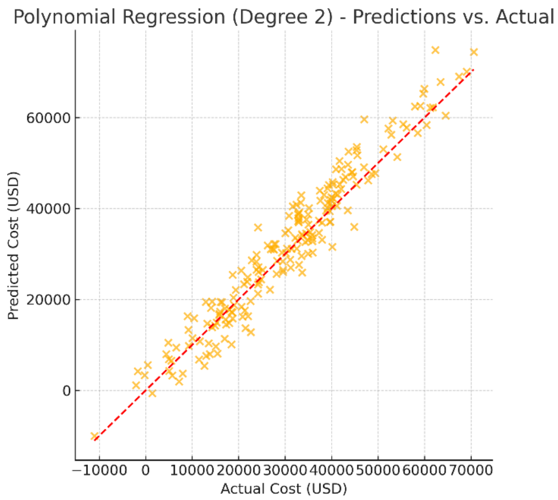
This section examines comparisons between predicted and actual values for different regression models for predicting hospital costs, emphasizing the accuracy and reliability with which the regression models predict hospital costs. The evaluation process concentrates on the effectiveness of the proposed optimized Gradient-boosting methodology in decreasing deviations and improving predictive performance.

Figure 5 shows the predicted vs actual values for various regression models, and each plot shows the accuracy of the respective method. The pattern of multiple linear regression around the diagonal indicates that it is a poor predictor of the dependents. A tighter clustering along the diagonal for Polynomial Regression (Degree 2) indicates improved performance via non-linear modelling. While SVM Regression can capture patterns, significant deviations exist for higher actual values than predicted.

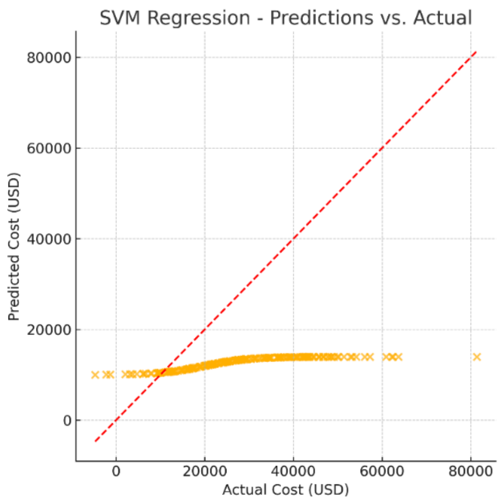
Compared with the diagonal, we find that ensemble methods, such as Random Forest(Default), provide better alignment, i.e., more accurate predictive accuracy. XGBoost (Default) further refines this alignment by closely predicting actual values. The Proposed Optimized XGBoost model also has a tight clustering along the diagonal line, which indicates the least deviation and almost the best predictive performance. It demonstrates that the optimization process effectively helps it achieve hospital cost forecasting.


Multiple Linear Regression - Predictions vs. Actual
Predicted Cost (USD)
Actual Cost (USD)

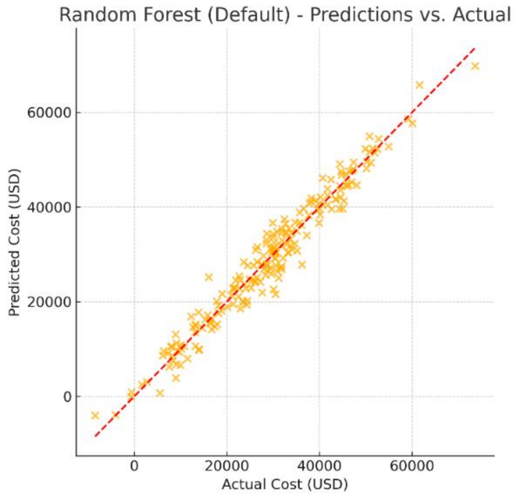
(a)


Polynomial Regression (Degree 2) - Predictions vs. Actual
Predicted Cost (USD)
Actual Cost (USD)

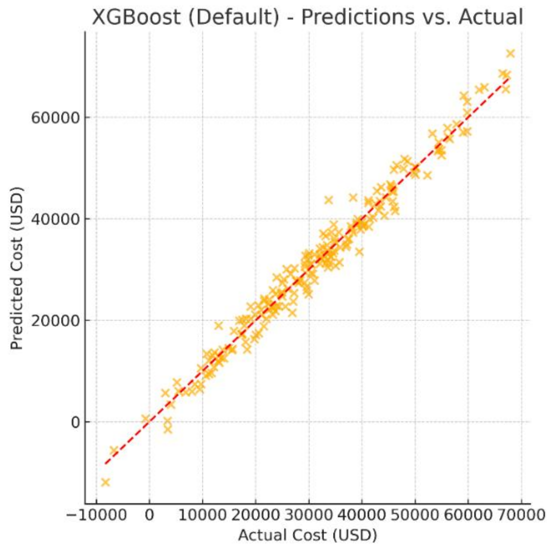
(b)


SVM Regression - Predictions vs. Actual
Predicted Cost (USD)
Actual Cost (USD)

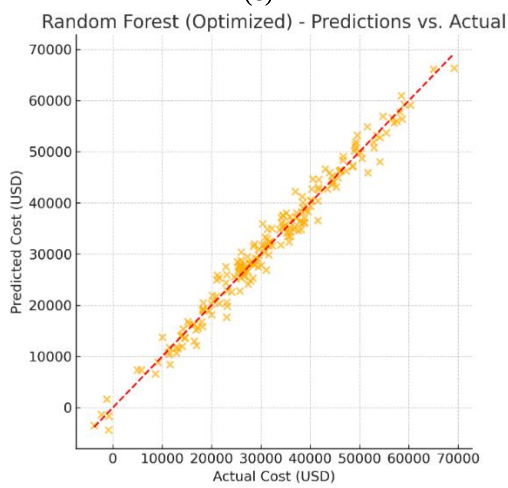
(c)


Random Forest (Default) - Predictions vs. Actual
Predicted Cost (USD)
Actual Cost (USD)

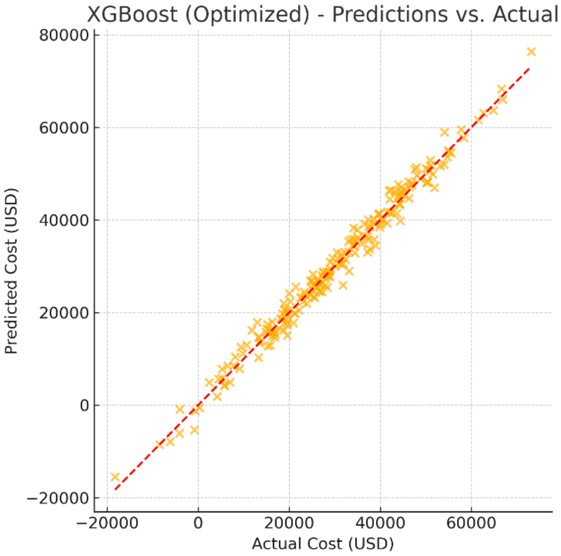
(d)


XGBoost (Default) - Predictions vs. Actual
Predicted Cost (USD)
Actual Cost (USD)

(e)


Random Forest (Optimized) - Predictions vs. Actual
Predicted Cost (USD)
Actual Cost (USD)

(f)



(g)

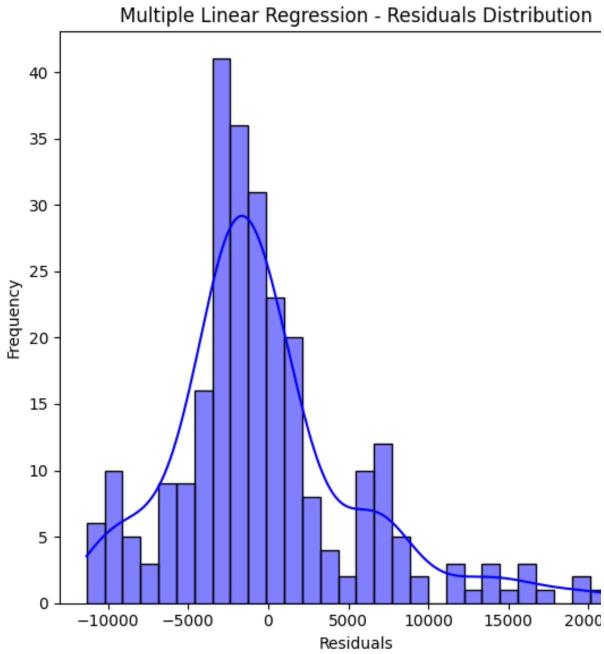
Figure 5: Predictions vs. actual cost (USD) for (a) Multiple Linear Regression, (b) Polynomial Regression (Degree 2), (c) SVM Regression, (d) Random Forest (Default), (e) Random Forest (Default), (f) XGBoost (Optimized), and the (g) Proposed Optimized XGBoost. The Proposed Optimized XGBoost demonstrates the best alignment with the diagonal, reflecting the highest predictive accuracy and minimal deviations among all models.

5.2 Residual distribution of the proposed framework models

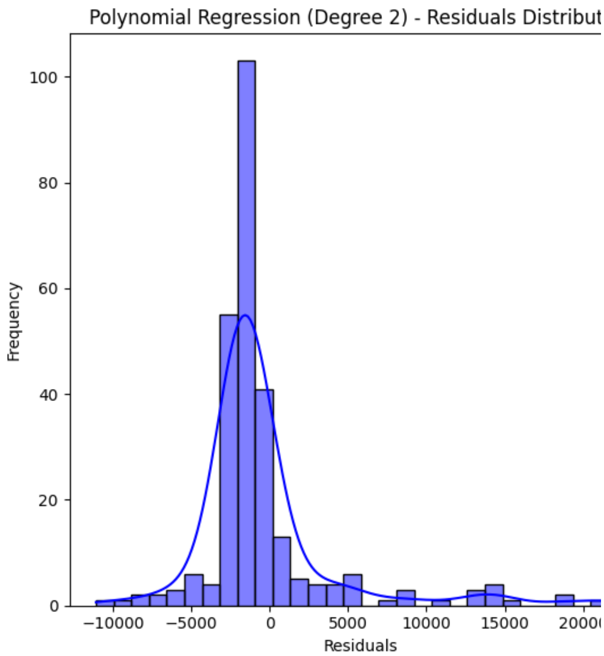
This section then analyzes the residual distribution of different regression models applied in the presented framework to understand errors in prediction and patterns. This paper uses the Proposed Optimized Gradient Boosting Methodology to analyze the models' error reduction capacity. Figure 6. shows residuals around zero (i.e., how accurate and biased the models are).

Different regression models are compared against each other through the distribution graphs of residuals. Residuals of Multiple Linear Regression show a wider spread, indicating more significant prediction errors. Polynomial Regression (Degree 2) reduces the spread, reflecting its ability to model non-linear relationships. Looking at the distribution of SVM Regression, we observe a skewed distribution along with significant outliers, implying that SVM Regression might not be able to handle complicated relations.

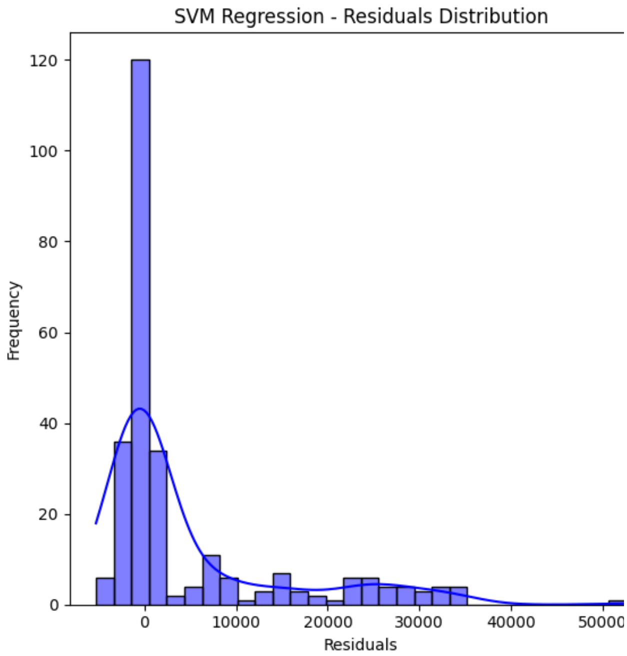
Improved accuracy is found in Random Forest (Default), with a more concentrated residual distribution. Further improvement is made by XGBoost (Default), with most residuals close to zero. Finally, the Proposed Optimized Gradient Boosting Methodology (Optimized XGBoost) had the narrowest spread and the lowest spread around the mean, indicating little error and good predictive accuracy out of all models. Thus, optimization has been shown to reduce prediction errors and improve model performance.



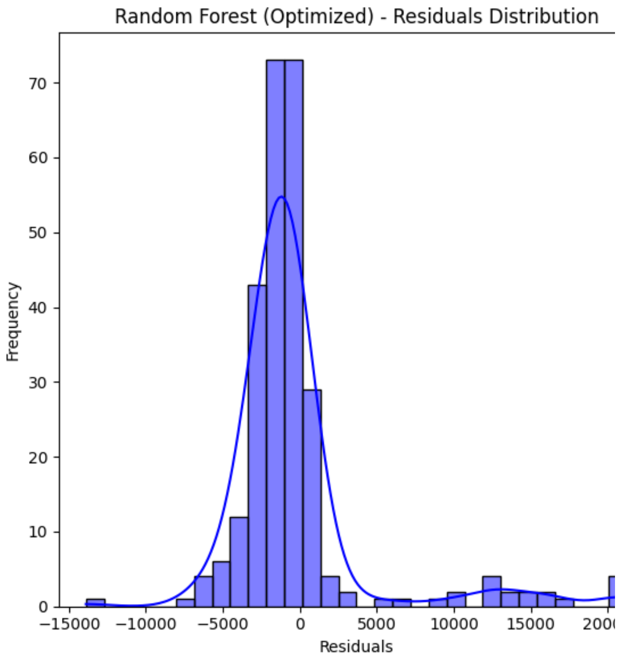
(a)



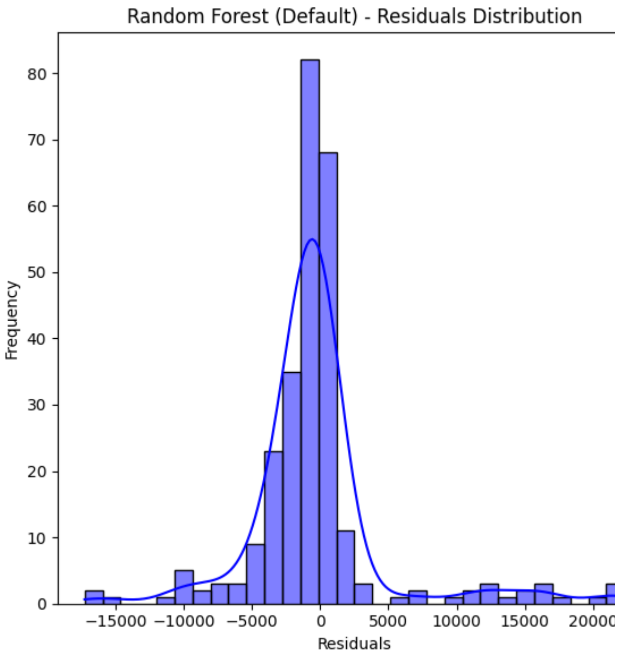
(b)


SVM Regression - Residuals Distribution
Frequency
Residuals

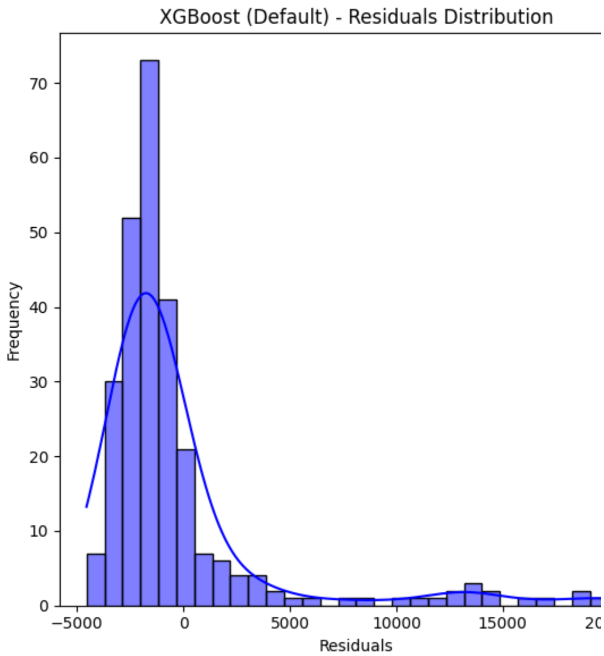
(c)


Random Forest (Optimized) - Residuals Distribution
Frequency
Residuals

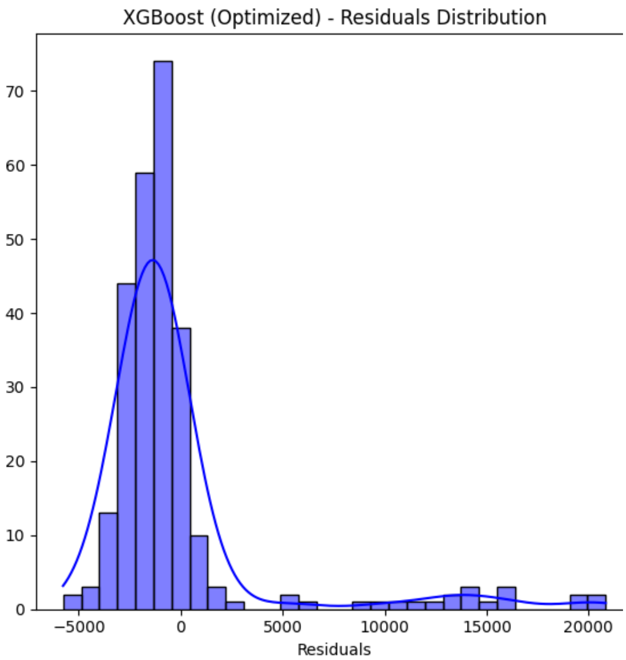
(e)


Random Forest (Default) - Residuals Distribution
Frequency
Residuals

(d)


XGBoost (Default) - Residuals Distribution
Frequency
Residuals

(f)



(g)

Figure 6: Residuals distribution of various regression models Proposed framework, including Multiple (a) Linear Regression, (b) Polynomial Regression (Degree 2), (c) SVM Regression, (d) Random Forest (Default), (e), Random Forest (Optimized), (f) XGBoost (Default), and the (g) Proposed Optimized XGBoost, showcasing progressive improvements in error reduction, with the Proposed Optimized XGBoost achieving the most symmetric and narrowest residual distribution, reflecting superior predictive accuracy.

5.3 Comparison proposed framework models

The values of R2 presented in Table 5 and Figure 7 show the additional performance enhancements when non-linear modelling and hyperparameter optimization are done. Multiple Linear Regression returned the lowest R^2 score (0.73) expected, as its ability to solve such complex, non-linear relationships in healthcare cost data is limited. This performance was modestly improved (R^2 = 0.79) using non-linearity in the Polynomial Regression, though global polynomial assumptions still constrained it.

The R^2 of the SVM Regression model is 0.81, which outperforms the models we tested based on the outcome of the dataset but underperformed when tested on other data sets. All models (R^2 = 0.85) were superseded by Random Forest (Default) and XGBoost (Default) performing (R^2 = 0.85 and 0.86, respectively), and this outcome is attributable to the capacity of Random Forest (Default) and XGBoost (Default) to discover feature interactions and different responses.

Optimized Random Forest brought the most gains, as it produced an R^2 of 0.88, while Optimized XGBoost was slightly better than it with R^2 of 0.89. Although the performance gap between the two optimized ensemble models is on the order of 0.01, the resulting difference and the decreased residual error (as shown in Fig. 5) indicate an advantage of gradient boosting's sequential error correction. The slight difference suggests that both methods are viable, and perhaps the final compromise would be performance, training efficiency, interpretability, etc.

Table 5: Comparing performance metrics (R^2 values) of the proposed framework with other models, including Multiple Linear Regression, Polynomial Regression, SVM Regression, Random Forest (Default), and XGBoost (Default)

Model	**R² Value**
Multiple Linear Regression	0.73
Polynomial Regression	0.79
SVM Regression	0.81
Random Forest (Default)	0.85
XGBoost (Default)	0.86
Random Forest (Optimized)	0.88
Optimized XGBoost	**0.89**

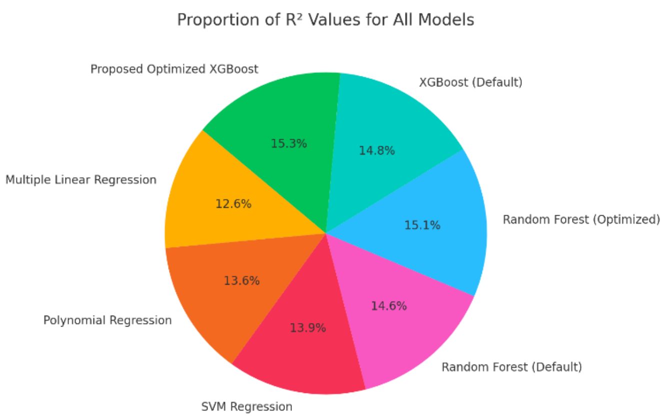


Figure 7: Distribution of R^2 values throughout the proposed framework made proportionally. For each model, the relative predictive accuracy is illustrated, and the Proposed Optimized XGBoost achieves the best performance, after which the following model, Random Forest (Optimized), comes next.

The Optimized Gradient Boosting Methodology enhances performance in predicting public hospital costs. The comparison of R^2 .In scores across several models, the implications of using advanced ensemble techniques and optimization strategies are demonstrated. Following that, we ran Multiple Linear Regression, achieving an R^2. With a score of 0.73, it is limited in finding complex factors driving healthcare expenditure data since it assumes relations are linear. The addition of model flexibility through Polynomial Regression increased performance to 0.79, partly due to improved accuracy by exploiting the non-linear relationship between the variables.

Additional machine learning models made further advances in predictive accuracy. SVM Regression achieved an R^2. It benefits from its ability to model non-linear patterns and has a score of 0.81. These traditional approaches were outpaced by Ensemble methods, with Random Forest (Default) showing a very impressive R^2. The strength of ensemble learning on feature interaction is reflected in the score of 0.85. The Random Forest model was further optimized R^2. It proves the value of improving predictive performance with hyperparameter tuning, reducing this score to 0.88.

The XGBoost (Default) slightly outran the default Random Forest with an R^2 . The gradient boosting framework achieves superior accuracy at a score of 0.86. However, the Optimized XGBoost achieved the highest performance and was able to deliver an R^2 With a score of 0.89, this also becomes the best-performing model. We attribute this improvement to advanced hyperparameter optimization, which improves things like learning rate, tree depth, and regularization parameters, which makes the model better to generalize.

These results can have substantial implications for hospitals' budgeting. In particular, the optimized ensemble models show very high predictive accuracy, which renders them excellent intelligent cost accounting and resource allocation tools. Our results show the criticality of model optimization and the capability of gradient-boosting algorithms to deal with complex, non-linear relationships in healthcare data. Finally, the Proposed Optimized Gradient Boosting Methodology introduces a revolutionary approach to deploying data to optimize financials and make intelligent choices so that public hospitals can use this methodology at their doorsteps.

The XGBoost model is applied over boosting rounds, Random Forest is applied over the number of trees, and training and testing R^2 scores for both are shown in Figure 8. With XGBoost (Figure 8a), we get a rapid boost in R^2 until about 100 iterations; after that, everything levels off, so we can say that it converged. The R^2 curves can be slightly different after 200 rounds between training and testing, which implies minor overfitting, but this was overcome by applying early stopping. The training and testing curves from Random Forest (Figure 8b) stop stabilizing after about 80 trees. It may be overfitted due to a randomized construction of the trees and the use of the regularization parameters (e.g., min_samples_split). We know these trends support ensemble methods' robustness and stability as much as the tuned XGBoost.

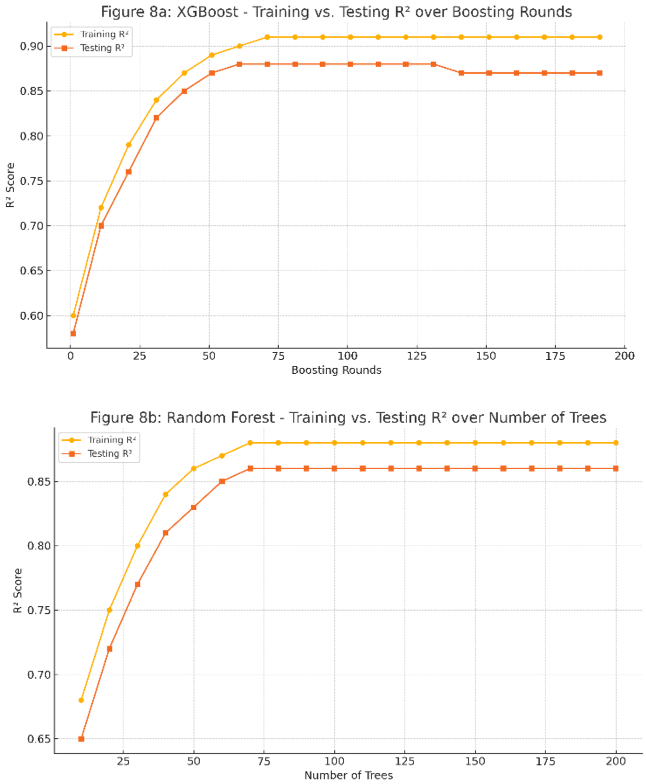
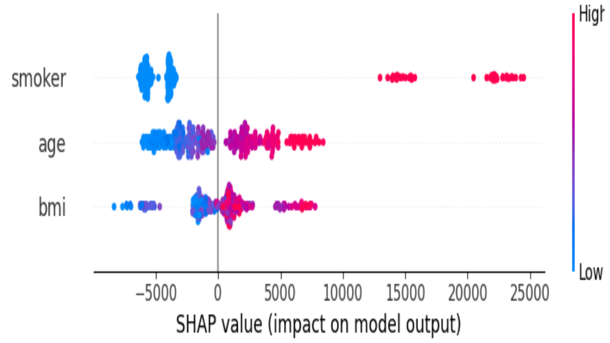


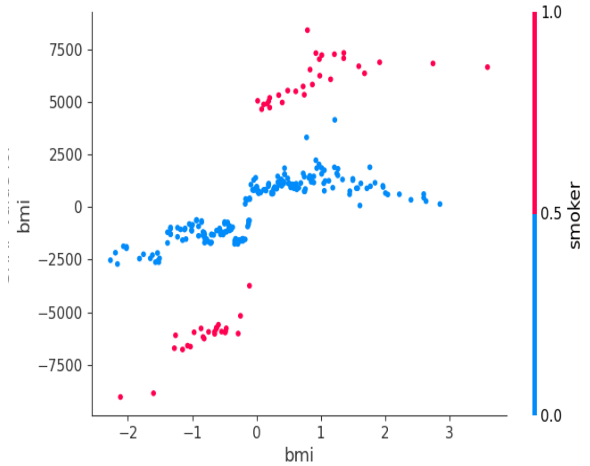
Figure 8: Training vs Testing R^2 scores for the optimized models. Figure 8a shows the XGBoost model converging after ~150 boosting rounds with slight overfitting mitigated by early stopping. Figure 8b displays the Random Forest model stabilizing after ~80 trees, indicating minimal overfitting and robust generalization.

5.4 Interpretability with SHAP analysis

In Figure 9, three SHAP visualizations provide the interpretability of the optimized XGBoost model. The SHAP summary plot, which figures out the essential features of the model by sorting them out based on how much they contribute to the model output in its training data (shown in Figure 9a), shows features by their impact on the model's production in the dataset. It is not an absolute contribution value but shows the direction and relative magnitude of influence a feature has on a prediction. An advantage of using SHAP values for identifying features that tend to affect cost predictions heavily is that SHAP values for features such as smoking status, BMI, and age contain the highest lowering and raising ranges, indicating these features often lead to positive or negative effects on cost predictions. Figure 9b displays the graph of BMI as a model input with a non-linear relationship, meaning that lower values don't influence that much, while higher values increase cost. As mentioned above, the plot also depicts an interaction with smoking status based on color encoding. A smoking status dependency plot (Figure 9) shows that those with SHAP values that tend to increase predicted costs in the sense they are more likely to be smokers in the real world tend to be higher. These plots point out that while smoking status certainly has a strong directionality, BMI and age have less sharp and more gradual, non-linear increases, pointing out the model's power to handle more complex cost-driving patterns.



(a)



(b)

Figure 9: SHAP analysis results of the Proposed Optimized Gradient Boosting Methodology. Similarly, at the top (a) is the SHAP summary plot indicating the impact on cost prediction by the features. Our SHAP dependency plot (b) displays the non-linear relation between BMI and costs and the interaction of the smoker feature.

6 Discussion

This research finds the potential of artificial intelligence and machine learning in transforming public hospital budgeting. The Optimized Gradient Boosting Method outperforms the predictive results, although the optimized XGBoost model obtains an R^2.

It involves running several regression analyses and looking at our results: Our optimized ensemble models (Random Forest and XGBoost) exhibit better predictive accuracy than traditional regression approaches, and the XGBoost model is at R^2 of 0.89. Compared to other machine learning methods that are tested (e.g., SVM and standard ensemble models), this is a modest improvement. Still, it does provide proof of the principle that hyperparameter searching and ensemble models have the potential to capture implicit, complex relationships between cost drivers. It is consistent with what is known about healthcare cost modeling in the literature, where often ensemble learning is preferred due to its ability to handle non-linearity and feature interactions. Nevertheless, these performance differences in this study were incremental, and the model selection must also include interpretability, computational efficiency, and implementation context.

In addition, this study is essential for integrating SHAP analysis, which stains transparency by attributing model predictions to each feature on a per-patient basis. Since the interpretability of this model means that decision-makers can now understand how such variables as smoking status or BMI affect cost prediction, they will be able to buy plans more easily. SHAP analysis showed smoking status, BMI, and age to be the most impactful factors in healthcare costs. The findings provide information for designing targeted health interventions (e.g., smoking cessation programs), deciding where to allocate scarce resources to the high-risk demographic, and how to stratify the costs for high-risk populations, making the insights useful for budget planning and designing policy. This interpretation ensures accuracy and transparency, builds trust among the stakeholders, and guides data-driven policy development. For example, knowing that smoker status is a very significant cost driver can help design targeted smoking cessation programs that are consistent with public health objectives and financial objectives.

This framework goes beyond forecasting individual healthcare costs, and this can also be applied to policy interventions on islands that consume cost drivers as identified by the model. For instance, SHAP analysis insights may help hospitals detect that smoking status is a significant cost driver and, based on this, develop or target a particular smoking cessation program to high-cost patient segments. In addition, the model applies to flag patients with an elevated BMI to help initiate prevention treatments aimed at reducing obesity-related complications. On the hospital finance team's administrative level, risk-adjusted budget allocation strategies can be designed according to region, age group, or behaviour risk factor that aligns with the anticipated cost impact. It can be done during implementation through integration with existing hospital-ERP systems or as a part of the BI dashboard, which will periodically train on updated patient data.

Our optimized framework is competitive (and in many cases better) relative to prior studies using machine learning to predict healthcare costs. It is shown in Table 1 (from Literature Review) that existing models such as Random Forest, gradient boosting or other commonly employed methods often report R^2 values between 0.78

and 0.86 on real-world healthcare datasets (e.g., Vimont et al., Mazumdar et al.). Our optimized XGBoost model, however, showed the R^2 of 0.89, MAE of 2502.36 and minimal residual variance among tested models. This improvement can be attributed to deep hyperparameter optimization, the addition of domain-relevant features, and the importance of calibration of the SHAP feature.

From a residual analysis viewpoint, our framework has a tight-centered error distribution with minimal skewness, indicating robust generalization. Traditional models such as linear regression, when used to have its residual distribution, had wider variance, especially at the higher cost levels—at this point, these models proved unable to capture the non-linear interactions every day in healthcare spending patterns.

It is partly because our model can balance model complexity with interpretability. Using SHAP analysis, we increased the prediction's transparency—how the model predicts—and verified the feature importance rankings and validated them with empirical and domain-specific evidence.

SHAP analysis showed that the most influential predictors of healthcare costs were smoking status, BMI, and age. These findings agree with what is already known about public health.

- SHAP values of the effects proved to be the most critical factor on cost, with smoking status showing the most substantial positive influence. It is to be expected since smoking is known to be a well-known risk factor for chronic diseases such as cardiovascular and respiratory conditions, which substantially increased healthcare utilization. Additionally, this variable had a binary nature that led to model clarity and decision boundaries.
- BMI is important because it is a surrogate for obesity-related complications, such as diabetes and orthopedic conditions. In SHAP, dependency plots exhibited a non-linear, threshold-based cost escalation for BMI > 30, as would be expected given obesity classification thresholds.
- There was a moderately and steadily increasing cost as with age. On the other hand, while the effect of age was linear, unlike smoking or BMI, the model could estimate it using only linear interaction terms rather than complex interaction terms.

There were other features — such as number of children, region, and sex — that had comparatively less impact. Still, scattered SHAP values showed weaker or inconsistent influence on the predicted cost outcomes.

This study is aware of some of its limitations despite its successes. As with the availability of high-quality, comprehensive datasets, the accuracy and scalability of the framework require a solution to some of the issues. Ensemble models can also impose computational demand, especially for optimization, with potentially prohibitive costs in resource-constrained settings. In addition to natural world hospital systems, such frameworks must be integrated while addressing data privacy requirements and making ethical decisions transparent.

Scalability: The manuscript acknowledges scalability as a challenge, but numerous strategies exist to scale beyond the experiment. If you have large datasets or a multi-hospital system, distributed computing frameworks like Apache Spark or Dask can run distributed data prep by dividing up pieces of your data or models to train them in parallel. In addition, the proposed framework is also fully compatible with cloud-based ML platforms (e.g., AWS SageMaker, Google Vertex AI), which offer auto-scaling infrastructure and a managed environment for model deployment. In future work, federated learning techniques may be explored to support multiple institutions such that model training is feasible from decentralized data silos without sacrificing privacy. Such strategies make it possible to keep the framework up-to-date and productive with the institution's growing data volume and size.

Finally, this work provides a firm foundation to utilize machine learning to optimize public hospital budgeting. The Proposed Optimized Gradient Boosting Methodology is a hybrid methodology of predictive accuracy, interpretability, and practicality to create a robust architecture of intelligent cost accounting and financial optimization. Avenues for future work include expansion of coverage in the scope of the dataset, scalability, and evaluation of the framework's impact in the real world in the hospital setting. If this proposed methodology can tackle these challenges, it might revolutionize financial management in public healthcare by improving resource allocation efficiency and patient care outcomes.

Practical Implications: The proposed framework provides a reasonable basis for implementation into actual hospital systems, requiring only modest infrastructures such as mid-range dedicated servers, cloud-based platforms, or standard Python environments, such as the Docker-provided ones. All tools (XGBoost and SHAP) are open source and do not require additional software. Data preparation, model training for the first time, and short workshops for staff are the most critical implementation costs, and the initial costs are estimated between $15,000 and $50,000, depending on the hospital scale. It is flexible enough to be compatible with secure on-premise systems to address data privacy concerns. It can be brought into these existing ERP or BI systems for risk-adjusted budgeting and policy decisions. It is a sustainable data-driven hospital financial management tool because it is adaptive to local data sets and scalable through periodic retraining.

Ethics and Privacy: This study is bound to observe data privacy and ethical standards rigidly. All the data that we have used in our dataset was fully anonymized, with no personally identifiable information in it at all. However, in real-world hospital deployments, they must implement the necessary data governance framework according to HIPAA, GDPR, or other local regulations. The list of things included in this is data encryption, access controls, secure storage, and role-based permissions. Secondly, the proposed framework is compatible with institutions' existing infrastructure. It retains its control over patient data since institutions do not have to give up control over their patient data by sharing it with third-party clouds. In addition to providing transparency in model output interpretability and guaranteeing ethical AI deployment in health care, SHAP analysis integration is also addressed.

7 Conclusion

This thesis proposes an interpretable machine learning framework for public hospital budgeting that balances an interpretable model and budgeting practicality through a combination of predictive model explanations. The framework optimizes ensemble methods, of which Random Forest & XGBoost are good examples, along with SHAP analysis to obtain correct forecasting of healthcare costs and the ability to clarify key cost drivers (e.g. smoking status, BMI, etc) about the patients. The term' hybrid' reflects the integration of model performance with stakeholder-relevant interpretability. In contrast, the framework's architecture comprises modular steps such as data preprocessing, model training, hyperparameter tuning, and SHAP interpretation. Practicality is achieved by practising the use of open source tools, minimal infrastructure requirements, and compatibility of infrastructure with existing hospital IT. While the performance gains over baseline models are small, the interpretability and operational relevance of the framework enables its use to guide focused interventions and resource planning. For instance, hospitals could apply the insights to back preventive care efforts for high-risk populations or allocate funds according to the risk-adjusted budget. Future work will investigate distributed learning approaches that would scale the framework to a larger, multi-hospital dataset and apply an existing policy impact method to real-world policy impacts in healthcare settings. However, further studies are needed to evaluate long-term outcomes and improve Financial Management decisions in the hospital regarding cost forecasting using Explainable AI together with ensemble learning.

References

[1] A. Shiwlani, M. Khan, A. M. K. Sherani, M. U. Qayyum, and H. K. Hussain, "REVOLUTIONIZING HEALTHCARE: THE IMPACT OF ARTIFICIAL INTELLIGENCE ON PATIENT CARE, DIAGNOSIS, AND TREATMENT," *JURIHUM: Jurnal Inovasi dan Humaniora*, vol. 1, no. 5, pp. 779-790, 2024. Retrieved from https://jurnalmahasiswa.com/index.php/jurihum

[2] K. J. Prabhod, "The Role of Artificial Intelligence in Reducing Healthcare Costs and Improving Operational Efficiency," *Quarterly Journal of Emerging Technologies and Innovations*, vol. 9, no. 2, pp. 47-59, 2024. https://doi.org/10.58532/nbennurch302

[3] D. Brunner, C. Legat, and U. Seebacher, "Towards Next Generation Data-Driven Management," *Collective Intelligence: The Rise of Swarm Systems and their Impact on Society*, p. 152, 2024. https://doi.org/10.1201/9781032690711-8

[4] N. A. Wani, R. Kumar, J. Bedi, and I. Rida, "Explainable AI-driven IoMT fusion: Unravelling techniques, opportunities, and challenges with Explainable AI in healthcare," *Information Fusion*, p. 102472, 2024. https://doi.org/10.1016/j.inffus.2024.102472

[5] A. Vimont, H. Leleu, and I. Durand-Zaleski, "Machine learning versus regression modelling in predicting individual healthcare costs from a representative sample of the nationwide claims database in France," *The European Journal of Health Economics*, vol. 23, no. 2, pp. 211-223, 2022. https://doi.org/10.1007/s10198-021-01363-4

[6] M. Mazumdar *et al.*, "Comparison of statistical and machine learning models for healthcare cost data: a simulation study motivated by Oncology Care Model (OCM) data," *BMC health services research*, vol. 20, pp. 1-12, 2020. https://doi.org/10.1186/s12913-020-05148-y

[7] B. Langenberger, T. Schulte, and O. Groene, "The application of machine learning to predict high-cost patients: A performance-comparison of different models using healthcare claims data," *PloS one*, vol. 18, no. 1, p. e0279540, 2023. https://doi.org/10.1371/journal.pone.0279540

[8] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5-32, 2001. https://doi.org/10.1023/a:1010933404324

[9] S. Ramraj, N. Uzir, R. Sunil, and S. Banerjee, "Experimenting XGBoost algorithm for prediction and classification of different datasets," *International Journal of Control Theory and Applications*, vol. 9, no. 40, pp. 651-662, 2016. https://doi.org/10.7717/peerj-cs.2451/fig-12

[10] S. Nanglia, M. Ahmad, F. A. Khan, and N. Jhanjhi, "An enhanced Predictive heterogeneous ensemble model for breast cancer prediction," *Biomedical Signal Processing and Control*, vol. 72, p. 103279, 2022. https://doi.org/10.1016/j.bspc.2021.103279

[11] J. Abdollahi, B. Nouri-Moghaddam, and M. Ghazanfari, "Deep Neural Network Based Ensemble learning Algorithms for the healthcare system (diagnosis of chronic diseases)," *arXiv preprint arXiv:2103.08182*, 2021. https://doi.org/10.1007/s00521-021-06459-9

[12] H. Kwon, J. Park, and Y. Lee, "Stacking ensemble technique for classifying breast cancer," *Healthcare informatics research*, vol. 25, no. 4, pp. 283-288, 2019. https://doi.org/10.4258/hir.2019.25.4.283

[13] F. Ali *et al.*, "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion," *Information Fusion*, vol.

63, pp. 208-222, 2020. https://doi.org/10.1016/j.inffus.2020.06.008

[14] D. D. Rufo, T. G. Debelee, A. Ibenthal, and W. G. Negera, "Diagnosis of diabetes mellitus using gradient boosting machine (LightGBM)," *Diagnostics*, vol. 11, no. 9, p. 1714, 2021. https://doi.org/10.3390/diagnostics11091714

[15] A. Y. Krishna, K. R. Kiran, N. R. Sai, A. Sharma, S. P. Praveen, and J. Pandey, "Ant Colony Optimized XGBoost for Early Diabetes Detection: A Hybrid Approach in Machine Learning," *Journal of Intelligent Systems & Internet of Things,* vol. 10, no. 2, 2023. https://doi.org/10.54216/jisiot.100207

[16] K. Amarasinghe, K. T. Rodolfa, H. Lamba, and R. Ghani, "Explainable machine learning for public policy: Use cases, gaps, and research directions," *Data & Policy,* vol. 5, p. e5, 2023. https://doi.org/10.1017/dap.2023.2

[17] A. Tursunalieva, D. L. Alexander, R. Dunne, J. Li, L. Riera, and Y. Zhao, "Making Sense of Machine Learning: A Review of Interpretation Techniques and Their Applications," *Applied Sciences,* vol. 14, no. 2, p. 496, 2024. https://doi.org/10.3390/app14020496

[18] M. Van der Schaar *et al.*, "How artificial intelligence and machine learning can help healthcare systems respond to COVID-19," *Machine Learning,* vol. 110, pp. 1-14, 2021. https://doi.org/10.1007/s10994-020-05928-x

[19] W. Ding, M. Abdel-Basset, H. Hawash, and A. M. Ali, "Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey," *Information Sciences*, vol. 615, pp. 238-292, 2022. https://doi.org/10.1016/j.ins.2022.10.013

[20] N. Rane, S. Choudhary, and J. Rane, "Explainable Artificial Intelligence (XAI) in healthcare: Interpretable Models for Clinical Decision Support," *Available at SSRN 4637897*, 2023. https://doi.org/10.2139/ssrn.4637897

[21] M. Liu, Y. Ning, H. Yuan, M. E. H. Ong, and N. Liu, "Balanced background and explanation data are needed in explaining deep learning models with SHAP: An empirical study on clinical decision making," *arXiv preprint arXiv:2206.04050*, 2022. https://doi.org/10.3390/jrfm16040221

[22] M. A. Shakir *et al.*, "Developing Interpretable Models for Complex Decision-Making," in *2024 36th Conference of Open Innovations Association (FRUCT)*, 2024: IEEE, pp. 66-75. https://doi.org/10.23919/fruct64283.2024.10749922

[23] P. N. Srinivasu, N. Sandhya, R. H. Jhaveri, and R. Raut, "From blackbox to explainable AI in healthcare: existing tools and case studies," *Mobile Information Systems*, vol. 2022, no. 1, p. 8167821, 2022. https://doi.org/10.1155/2022/8167821

[24] S. Singhal, "Cost optimization and affordable health care using AI," *International Machine learning journal and Computer Engineering,* vol. 6, no. 6, pp. 1-12, 2023. https://doi.org/10.32996/jcsts.2023.5.4.23

[25] A. K. Leist *et al.*, "Mapping of machine learning approaches for description, prediction, and causal inference in the social and health sciences," *Science Advances,* vol. 8, no. 42, p. eabk1942, 2022. https://doi.org/10.1126/sciadv.abk1942

[26] J. Amann, "Machine learning in stroke medicine: Opportunities and challenges for risk prediction and prevention," *Artificial Intelligence in Brain and Mental Health: Philosophical, Ethical & Policy Issues*, pp. 57-71, 2022. https://doi.org/10.1007/978-3-030-74188-4_5

[27] M. Ordu, E. Demir, C. Tofallis, and M. M. Gunal, "A novel healthcare resource allocation decision support tool: A forecasting-simulation-optimization approach," *Journal of the operational research society,* vol. 72, no. 3, pp. 485-500, 2021. https://doi.org/10.1080/01605682.2019.1700186

[28] S. Joshi *et al.*, "Modeling conceptual framework for implementing barriers of AI in public healthcare for improving operational excellence: experiences from developing countries," *Sustainability,* vol. 14, no. 18, p. 11698, 2022. https://doi.org/10.3390/su141811698

[29] D. Patil, N. Rane, P. Desai, and J. Rane, "Machine learning and deep learning: Methods, techniques, applications, challenges, and future research opportunities," *Trustworthy Artificial Intelligence in Industry and Society*, pp. 28-81, 2024. https://doi.org/10.70593/978-81-981367-4-9_2

[30] J. Rane, S. Mallick, O. Kaya, and N. Rane, "Scalable and adaptive deep learning algorithms for large-scale machine learning systems," *Future Research Opportunities for Artificial Intelligence in Industry 4.0 and*, vol. 5, pp. 2-40, 2024. https://doi.org/10.70593/978-81-981271-0-5_2

[31] R. Ramya, S. Priya, P. Thamizhikkavi, and M. Anand, "The Pillars of AI Ethics: Transparency, Accountability, and Privacy," in *Responsible Implementations of Generative AI for Multidisciplinary Use*: IGI Global, 2025, pp. 85-110. https://doi.org/10.4018/979-8-3693-9173-0.ch004