Smart Guardian: An AI-Based Coronary Heart Disease Prediction System, Focusing on Your Cardiac Health!

Longdi Xian[a], Junhao Xu[b,*]

[a]Department of Psychiatry, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong
[b]Faculty of Computer Science and Technology, University Malaya, Malaysia

Abstract

Coronary heart disease (CHD) is a highly fatal disease. With the continuous advancement of scientific technology, various techniques such as machine learning, deep learning, Transformer, natural language processing, and artificial intelligence tools like GPT3.5, Chatglm, Gemini, and Baidu have been applied by many researchers to disease prediction. However, current research lacks people-centric systems in terms of algorithm accuracy and interpretability, which hinders the provision of intelligent decision-making and recommendations. Therefore, this study aims to construct a people-centric CHD prediction system. In this study, we utilize the dataset from the Behavioral Risk Factor Surveillance System (BRFSS) for experimentation. Advanced machine learning techniques such as XGBoost, LightGBM, RandomForest, as well as Transformer algorithms like FTtransformer and TabNet are evaluated using metrics like F1 and ACC to select high-performing models. During the model evaluation process, all factors influencing the CHD model are taken into account, and the most influential factors are further trained to construct a high-performance model. By integrating these retrained models, a comprehensive ensemble model for predicting CHD is formed. Simultaneously, following the principles of average risk value and comprehensive risk factors, after obtaining user input data, the ensemble model is used to determine the user's CHD risk percentage, relevant risk factors, and corresponding risk values. Important methods such as COT, Zero Shot, and word generation are employed to evaluate the prediction results by considering the outputs of intelligent tools such as GPT3.5, Chatglm, Gemini, and Baidu. Finally, utilizing GPT-3.5, personalized health guidance related to lifestyle and health factors is generated to promote individualized care and prevention. The experiments demonstrate that tree-based models and deep learning models exhibit similar performance characteristics. Therefore, the system integrates all the best models, including XGBoost, LightGBM, RandomForest, and TabNet, to obtain comprehensive prediction results. Additionally, the system supports multiple modes of operation, such as text input and voice input, and has provided reasonable health advice and guidance, meeting the initial expectations for accurate and people-centric CHD risk prediction.

Keywords: Coronary Heart Disease, CHD prediction, machine learning, deep learning, Transformer, human-centered system

1. Introduction

Coronary heart disease (CHD) stands as the primary cause of death and disability globally [1][2][3]. This condition arises when the arteries responsible for delivering blood to the heart constrict or become blocked due to the buildup of plaque [4][5]. Consequently, this leads to a reduction in the supply of blood and oxygen to the heart [6]. The automatic prediction of heart disease through contemporary methods currently represents one of the most crucial and challenging healthcare issues worldwide [7]. The World Health Organization (WHO) has reported that more than 17.9 million people across the globe succumb to cardiovascular diseases each year [8][9].

Currently, many scholars have dedicated significant research efforts to utilizing data mining techniques for extracting risk factors from unstructured text data [10]. Nazari et al. proposed a hybrid method based on fuzzy analytic hierarchy process and fuzzy inference system, and designed a clinical decision support system (CDSS), which is very useful for evaluating the possibility of heart disease [11]. The machine learning-based

*Corresponding author
Email address: `xujunhao1105@gmail.com` (Junhao Xu)

prediction system proposed by Yan et al. demonstrates strong discriminative capabilities in diagnosing coronary artery stenosis among suspected coronary heart disease patients. This system has the potential to reduce the frequency of unnecessary invasive examinations for patients who test negative and decrease the rate of missed diagnoses for patients who test positive [12].

Farman et al. present an intelligent medical monitoring system that has made significant breakthroughs in accuracy. However, since cardiology datasets typically contain both relevant features and a large number of redundant or irrelevant features, processing these features is not only time-consuming but also detrimental to classification accuracy. In this case, exploring better ways to eliminate irrelevant features and manage missing values effectively becomes a serious challenge [13].

In this Paper, a new intelligent medical monitoring system (*Human-Centered*[14]) is proposed that utilizes integrated machine learning and transformer technologies for heart disease prediction[15], ultimately providing specific medical recommendations. Firstly, machine learning and transformer technologies are employed to define the strong risk factor eigenvalues for CHD based on experimental datasets, and a system is developed based on these eigenvalues. The system utilizes multimodal methods to obtain and process user data input, including voice input, text input, expert diagnosis, etc. Secondly, feature fusion methods are applied to combine the data and store it within the medical database. Next, an integrated model (such as TabNet, FTTransformer, XGBoost, etc.) is used for prediction based on the obtained data, and a comprehensive judgment is made based on all prediction results to obtain the best result. Finally, the obtained risk results and the original user inputs are used to prompt using LLM and then connected to GPT3.5 and used Bert[16] to obtain the most accurate and self-compatible person-centered medical advice. The main contributions of this paper are as follows.

- A new framework for predicting heart disease has been proposed, incorporating an integrated multimodal approach for data input and an integrated model for multi-risk prediction, in line with a people-oriented concept.
- The model is built using a combination of multiple machine learning and transformer algorithms for feature selection, reducing the complexity and dimensionality of the dataset by eliminating irrelevant features. This process enhances the accuracy of feature acquisition.
- The latest GPT tools have been employed to provide intelligent medical advice at a personal level. This enables artificial intelligence to interact with humans and automatically derive the best personalized medical advice.

2. Literature Review

As technology continues to advance, numerous scholars have contributed greatly to the prediction of coronary heart disease. These include machine learning, deep learning and many other areas of research.

2.1. Coronary Heart Disease Prediction Using Machine Learning

In 2014, Meda & Bhogapathi concluded that Fuzzy Neural Genetic Algorithm (FNGA) and highlight subset selection could better predict heart disease [17], furthering the range of algorithms used for accuracy rates. Wan et al. used five different techniques in the WEKA tool to study the prediction of heart disease and came up with multi-layer perception neural network as the algorithm with the highest accuracy [18]. Fadnavis et al. again explored the accuracy of classification algorithms in data mining techniques and validated that the plain Bayesian classifier was the most accurate one based on Cleveland and Statlog heart disease databases [19]. In the same year, Rana et al. compared the accuracy of sequential minimal optimization (SMO), multilayer perceptron (MLP), random forest, and Bayes net for heart disease prediction based on the dataset collected from Ibn al-Bitar Hospital and the Baghdad Medical city. The highest accuracy of plain Bayesian was still obtained [20]. Wang et al. suggested that the sensitivity and accuracy of intracavitary electrocardiogram (IEGM) results based on data mining algorithm analysis were higher than that of conventional EGM [21], again validating the higher accuracy of data mining. Amanda H et al. used the South African Heart Disease dataset with 462 instances. They found that Naïve Bayes (NB) showed promise in detecting heart disease [22]. Daoud et al. found that LightGBM is significantly faster and more accurate than XGBoost for the same budget of hyperparameter optimization time [23]. Using a large number of hypertensive patients in Shenzhen, China, Zhenzhen Du et al. found that nonlinear models (K-nearest neighbor, random forest) outperformed linear models (logistic regression) for predicting heart disease using machine learning [24]. Rine Nakanishi et al. utilized a dataset of 66,636 asymptomatic subjects to compare machine learning of computed tomography (CT) and clinical variables for predicting death from cardiovascular diseases. Machine learning models outperformed traditional risk assessments [25]. Shen et al. proposed that machine learning methods, especially XGBoost, could effectively predict cardiovascular events during exercise assessment in coronary artery disease patients. However, they didn't construct or validate machine learning models [26]. Absar et al. concluded that the K-Nearest Neighbors (KNN) model in machine Learning was the most accurate for predicting heart disease [27], allowing the use of models in machine learning. Hassan et al. concluded that the random forest model (RF) in machine learning was more accurate than other gradient augmented trees (GBT) and multilayer perceptrons (MLP) [28]. Zhang et al., using the Z-Alizadeh Sani dataset and the LightGBM algorithm, conducted an experiment. The results show strong advantages in the models used for CAD detection [29]. A. Lakshmanarao et al. used a Kaggle public dataset and found that Random Forest and Extratree classifiers achieved the highest accuracy for predicting heart disease [30]. Yilmaz et al. used the IEEEDataPort database and concluded that the Random Forest (RF) model outperformed Support Vector Machine (SVM) and Logistic Regression (LR) models for heart disease prediction [31].

2.2. Coronary Heart Disease Prediction Using Deep Learning

Md Hassanuzzaman et al. utilized a transformer to build a model based on raw PCG data of 484 patients and obtained good results with an accuracy of 0.923 [32].Mandadi et al. discovered that accurately detecting coronary heart disease with Artificial Neural Networks (ANN) on the Cleveland dataset relies on network design, data quality, and preprocessing [33].Hyunok Yun et al. used data from participants in the Korean Genome and Epidemiology Study to employ a genetic risk score (GRS) significantly associated with predicting the incidence of coronary heart disease (CHD) in men and women. However, there was no significant association between the GRS and CHD [34].Kunlun Wang et al. used CTA data to predict coronary heart disease based on simple CNN, VGGNet, and ResNet, visual converters (ViT), and Vision Series Transformers (ViST), obtaining good statistical indicators for ViST [35].

2.3. Research Gap

Based on all the above literature review, most of the studies focused on specific techniques such as machine learning or deep learning. Future research could explore the integration of these techniques to create hybrid models that utilize the strengths of each method to improve the accuracy of coronary heart disease prediction. While some studies have reported high accuracy, the interpretability of the models is often not addressed. Given the critical nature of healthcare decision-making, research on interpretable AI techniques tailored to healthcare applications is critical, and this requires a combination of using state-of-the-art techniques (e.g., GPT3.5) to obtain understandable recommendations that provide actionable insights for both healthcare professionals and general users. Multiple studies have highlighted the importance of data quality and preprocessing, and future studies should explore standardized methods of data collection and preprocessing to ensure consistent and reliable results across different data sets.

3. Research Methodology

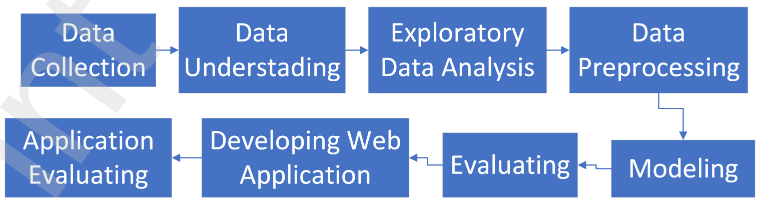


Figure 1: Research Design

This section discusses in detail the structure of the proposed research methodology and the basic structure of Towards a Human-Centered Medical System. The overall structure of this research methodology has been divided into several distinct stages to comprehensively present the information within each stage of the proposed research methodology. Finally, the fundamental framework for 'Towards a Patient-Centric Healthcare System' has been introduced, enabling users to predict patients' risk levels of heart disease based on their lifestyle and health factors and providing detailed, rational guidance using a Large Learning Model (LLM).

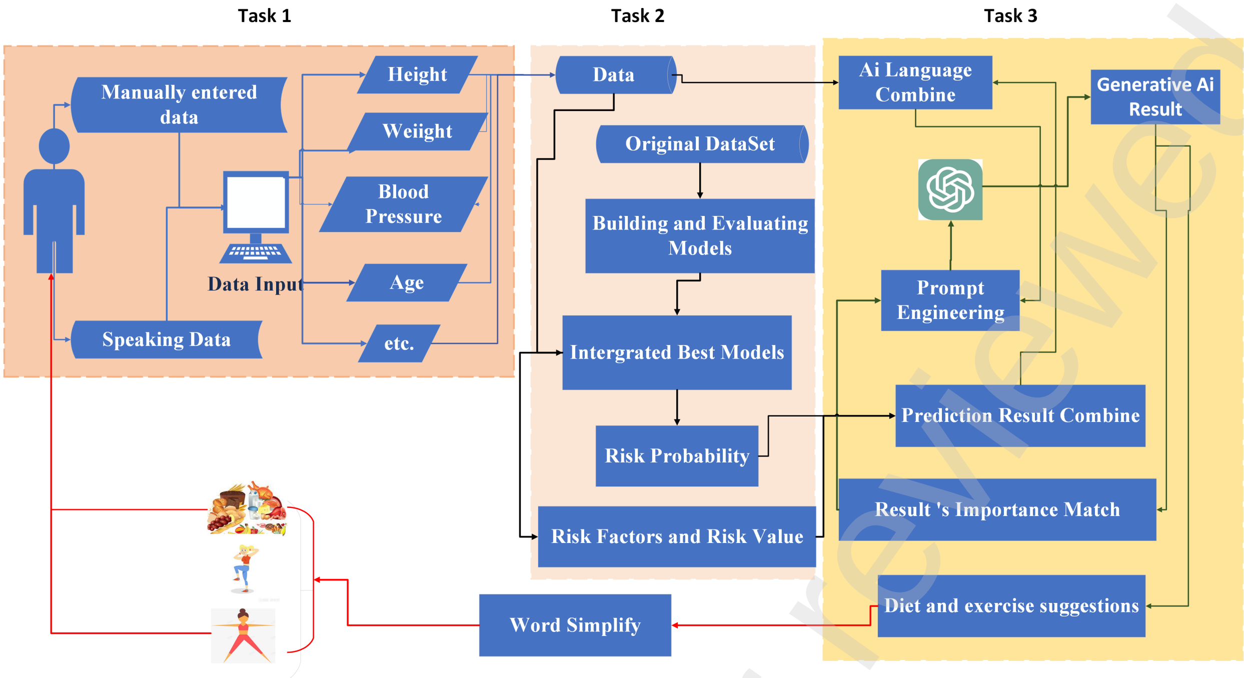


Figure 2: The basic structure of Towards a Human-Centered Medical System.

The architecture of the proposed research methodology is shown in Fig 1 and is divided into data collection, data comprehension, exploratory data analysis, and data preprocessing (here: cleaning of data, normalization of data types, definition of high-risk factors and re-generation of data, segmentation of data, and imbalance processing of data). Among these, the identification of high-risk factors and the training, prediction, and evaluation of the dataset are based on a selection of outstanding algorithms, including FTtransformer [36], XGBoost [37], TabNet [38], and others. The development of the web application is then carried out based on the best model obtained (Integrated use of the model) as well as the defined high-risk factors, and the details of the program are shown in Figure 3.

The basic structure of Towards a Human-Centered Medical System (HCMS) is illustrated in Figure 3. The HCMS has the main data from the user's real reconnaissance data, which comes from electronic medical devices (e.g., height, weight, etc.), assessments from specialists, as well as family history (e.g., stroke, mental illness, etc.) and internal judgments (e.g., hypertension, BMI, etc.). In order to provide convenience for people who can't write, we have set up multi-mode personal information input (text input, voice recognition input, e.g., Speech Recognition [39]). The purpose of HCMS is to predict a patient's disease risk based on real-time patient data. Therefore, we utilize health status as well as lifestyle habit prediction to predict heart disease based on the latest data collected.

After obtaining the data from the heart disease patients, the proposed system transmits the data to the relevant gateway device. This device integrates the data and stores it inside the medical database for further processing of the data (Figure 3, Task 1). The processed data are also predicted (Figure 3, Task 2) using the resulting best model (see Figure 4 for detailed steps). The next step is to use the data prediction results for further analysis and processing. This step includes data prediction, AI recommendation results, and human-centered result outputs (Figure 3, Task 3).

In the first step, the predicted results based on patient data are obtained, and in the second step, the results obtained from the prediction and the input data from the patient are fused into a statement recognizable by GPT3.5 based on the ordering of the importance of the features obtained in the previous method1, and then, the resulting spelled-out language is sent to the GPT3.5, and a preliminary artificial intelligence proposal is obtained using Prompt Engineering to obtain preliminary AI suggestions. The obtained recommendations are called again to the GPT3.5 to simplify and refine the obtained results to obtain the best human-centered recommendations. Ultimately, the best healthy diet and health (fitness duration, fitness style, etc.) recommendations are obtained based on the patient's own situation. As shown in Figure 3, the workflow of the proposed HCMS uses three successive layers: Data Input (Task 1), Data Collection and Prediction (Task 2), and Large Language Model (LLM) Proposal Suggestion (Task 3). These layers are specially discussed in the following subsections.

3.1. Data Input(Task 1)

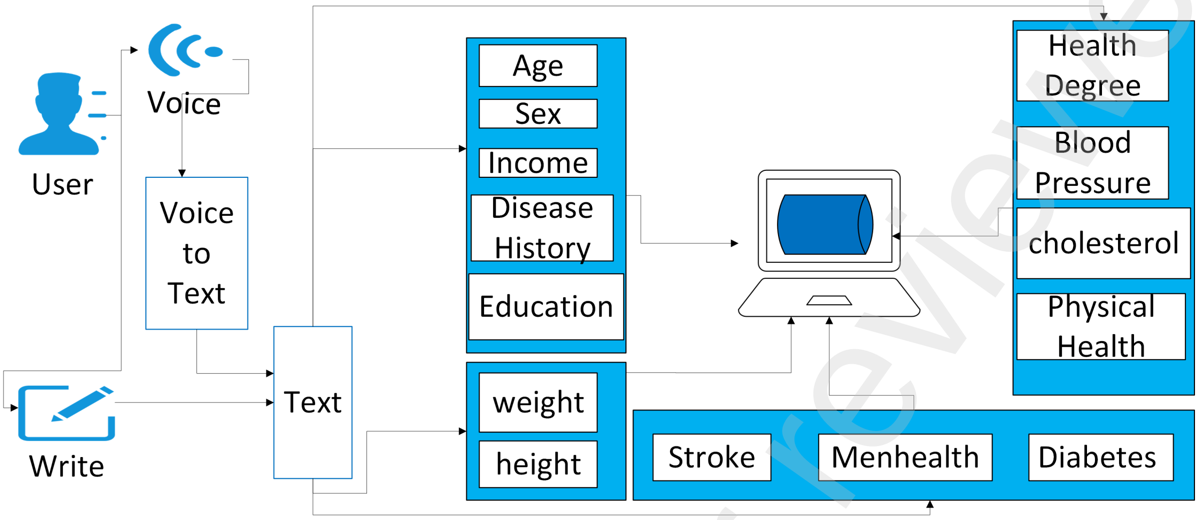


Figure 3: Data Input (Task1)

Given that the proposed system is human-centered, it comprehensively considers fundamental human data, such as age, gender, presence of diabetes, smoking history, and more, covering various age groups, genders, and individuals with disabilities. To ensure the system's applicability across diverse populations, we have implemented optimal response strategies for collecting data. Users can provide personal information through either voice or text input. In situations where users cannot read or may forget text, the system supports direct voice input, enhancing user convenience. For the specific implementation of voice input, we utilize the Speech Recognition technology. This technology is represented by the following formula:

$$\text{Voice Input} \xrightarrow[\text{(Speech Recognition)}]{\downarrow} \text{Text Representation} \tag{1}$$

This technology accurately transforms user-spoken information into a format that can be processed as text. The actual implementation of this process is not provided in specific code but is accomplished through the use of the Speech Recognition tool, ensuring both efficiency and accuracy in voice input. All collected personal health data will be securely stored in a database to ensure data integrity and privacy. The detailed operational steps of the system can be referenced in Figure 3.

3.2. Data Collection and Prediction(Task 2)

Figure 4 shows the training and evaluation of models, as well as specific details regarding the integration of models. For models, the new data is then stored within the medical database. Predictive model creation and adaptation are performed based on the collected dataset. This step utilizes the FTtransformer [36], XGBoost [37], TabNet [38], Random Forest [40], and other algorithms.

3.2.1. Data Processing

This dataset contains diverse data types, and for data type normalization, the `select_dtypes` method is initially used to choose specific data types, such as selecting columns of type 'object' as categorical features. For a specific column, like 'Age,' text categories are mapped to corresponding numerical values and converted to integers. Following this, the `LabelEncoder` is employed to label encode binary variables, and for non-binary variables, the `get_dummies` method from the `pandas` library is utilized for one-hot encoding. This preprocessing ensures that categorical features in the dataset are well-suited for machine learning model training.

Before constructing the Transformer model, the data underwent preprocessing with a primary focus on distinguishing between categorical and numerical features. `CATEGORICAL_FEATURES` include variables such as 'Diabetes', 'Stroke,' and others, while `NUMERIC_FEATURES` encompass 'BMI', 'DiffWalk,' and others. The

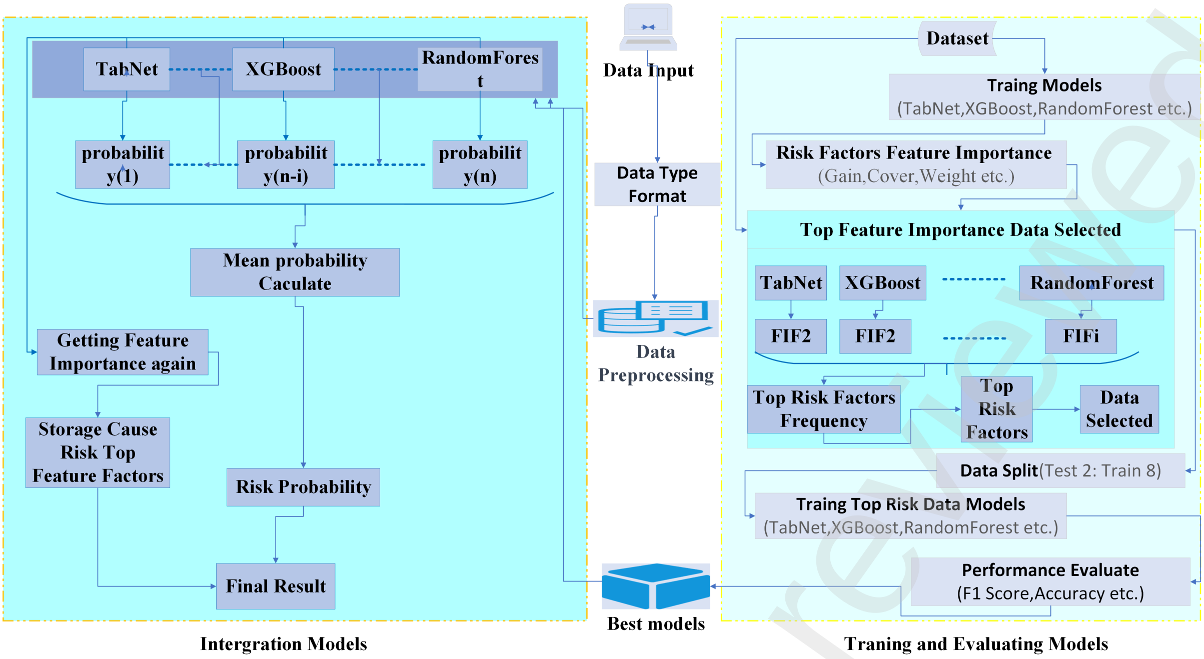


Figure 4: Prediction Models Work Principal

collective set of features is denoted as FEATURES, and the target feature is labeled as 'CHD'. The dataset will be divided into training (80%) and testing (20%) subsets using the `train_test_split` function from the `scikit-learn` library, with a specified random seed for reproducibility.

3.2.2. Modeling

This section describes the various experiments conducted to demonstrate the proposed HCMS for predicting heart disease and obtaining the corresponding recommendations. The preprocessing module analyzes the structured data for further processing. Additionally, the BRFSS publicly available dataset is utilized, and machine learning and deep learning are used to train the heart disease prediction model. Among the algorithms include FTtransformer[36], XGBoost[36], TabNet[36], RandomForest[36], AdaBoost[36], LightGBM[36]. During this period, the dataset was randomly divided into 70% and 30% for training and testing the above models, respectively, and 80% and 20% for training and testing, respectively.

In order to evaluate the proposed system framework based on multiple techniques, we have compared the accuracy of all the predictions made for feature selection in various ways. The detailed steps are shown in Figure 6, and the detailed computations are shown below.

3.2.2.1. Transformer Models

The construction of this model involves utilizing the Transformer model in TensorFlow for a binary classification task. The primary process of the model is outlined below:

Firstly, the data undergoes essential preprocessing steps. Numeric features are standardized using `StandardScaler` to ensure a mean of 0 and a standard deviation of 1. Subsequently, the `df_to_dataset` function is applied to transform the dataset into a TensorFlow dataset object for further processing. Next, the TabTransformer model is constructed using the `TabTransformer` class. This model is specifically designed for tabular data and is based on the Transformer architecture. In the model definition, hyperparameters such as numeric features, categorical features, encoding layers for categorical features, embedding dimensions, output dimensions, activation functions, etc., are specified. Refer to the 1 for details.

Specifically, the model structure is defined as follows: `TabTransformer(numerical_features, categorical_features, categorical_lookup)`. Where `numerical_features` and `categorical_features` are lists of numerical and categorical feature names, respectively, and `categorical_lookup` is a dictionary representing encoding layers for categorical features.

Subsequently, the model undergoes training. Initially, the `AdamW` optimizer is selected, which is an optimizer incorporating weight decay to adjust model parameters. `optimizer = tfa.optimizers.AdamW(learning_rate,`

Table 1: Table Transformer Hyperparameters

id	Hyperparameter	Value
1	embedding_dim	32
2	out_dim	1
3	out_activation	'sigmoid'
4	depth	4
5	heads	8
6	attn_dropout	0.2
7	ff_dropout	0.2
8	mlp_hidden_factors	[2, 4]
9	use_column_embedding	True
10	LEARNING_RATE	0.001
11	WEIGHT_DECAY	0.0001
12	NUM_EPOCHS	100
13	monitor	'val_output_loss'
14	mode	'min'
15	patience	20
16	restore_best_weights	True

`weight_decay)`. Where `learning_rate` is the step size at which the optimizer adjusts model parameters during training, and `weight_decay` is a regularization term that penalizes large weights to prevent overfitting.
Then, the model is compiled, setting the loss function to binary cross-entropy and the evaluation metric to accuracy. The Early Stopping callback is employed to monitor output loss on the validation set, preventing overfitting.
Finally, the model is trained using the `fit` method, taking in the training dataset and the validation dataset. After completing the model training, predictions are made using the trained model. The model is applied to the test dataset to generate predictions, and binary classification labels are produced based on the predicted probabilities.

3.2.2.2.Machine Learning Models

The process of constructing binary classifiers using XGBoost, AdaBoost, RandomForest, DecisionTree, and LightGBM involves several steps. Firstly, essential libraries, including pandas and the respective classifiers, are imported. Subsequently, classifier objects are initialized, and the `fit` method is employed to train the classifiers on the training set. Here is an overview of the working mechanisms of the core algorithms for each classifier.
XGBoost (eXtreme Gradient Boosting) is a machine learning algorithm based on Gradient Boosted Trees. It iteratively trains weak learners, typically decision trees, with each iteration attempting to correct the errors of the model from the previous iteration. The principle of XGBoost lies in adjusting model parameters by minimizing the gradient of the loss function, making each iteration more focused on correcting the errors from the previous one.

$$F_t(x) = F_{t-1}(x) + \gamma \cdot h_t(x) \tag{2}$$

Where: $F_t(x)$ is the prediction of the t^{th} tree,$F_{t-1}(x)$ is the cumulative prediction of the first $t-1$ trees, γ is the learning rate, controlling the contribution of each tree, $h_t(x)$ is the prediction correction of the t^{th} tree.

AdaBoost (Adaptive Boosting) is an ensemble learning algorithm that combines multiple weak learners to create a strong classifier. It assigns different weights to instances based on their classification errors, and each subsequent weak learner focuses on the misclassified instances. This adaptability improves overall accuracy, and the final prediction is a weighted sum of weak learner outputs. AdaBoost is advantageous for handling complex datasets and reducing bias and variance in the model.

$$W_i^{(1)} = N_1 \tag{3}$$

For the t^{th} iteration, the selection of weak learner $h_t(x)$ and the calculation of classification error:

$$\varepsilon_t = \sum_{i=1}^{N} w_i^{(t)} \cdot 1[h_t(x_i) \neq y_i] \tag{4}$$

Where ε_t is the classification error at the t^{th} iteration, $w_i^{(t)}$ is the weight of the i^{th} sample at the t^{th} iteration, and $1[h_t(x_i) \neq y_i]$ is the indicator function, equal to 1 if $h_t(x_i)$ (prediction of the i^{th} sample) is not equal to y_i (true label of the i^{th} sample).

Calculation of learner weight:

$$\alpha_t = \frac{1}{2} \ln \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right) \tag{5}$$

Where α_t is the weight assigned to the weak learner at the t^{th} iteration.

Update of sample weights:

$$w_i^{(t+1)} = w_i^{(t)} \cdot \exp\left(-\alpha_t \cdot y_i \cdot h_t(x_i)\right) \tag{6}$$

Where y_i is the true label of the i^{th} sample.

Combination of weak learners in the final strong learner:

$$H(x) = \text{sign}\left(\sum_{t=1}^{T} \alpha_t \cdot h_t(x) \right) \tag{7}$$

Where T is the number of iterations, and $H(x)$ represents the output of the final strong learner.

RandomForest is an ensemble learning algorithm that improves model performance by training multiple decision trees and combining their predictions. It reduces overfitting risk and enhances generalization by randomly selecting samples and features during training. Random Forest has advantages in handling high-dimensional data, managing missing values, and exhibiting robustness. Randomly select N samples from the dataset with replacement to form a bootstrap sample. Each sample has an equal probability of being selected: $D_1, D_2, \ldots, D_N$ Randomly select a subset of features (denoted as m) for each tree. This introduces diversity among the trees. For each bootstrap sample and feature subset, train a decision tree T_i independently: $T_1, T_2, \ldots, T_N$ For classification tasks, each tree provides a class prediction. The final prediction is determined by a majority vote.

$$\text{RandomForest}(x) = \text{mode}(T_1(x), T_2(x), \ldots, T_N(x)) \tag{8}$$

For regression tasks, each tree provides a numeric prediction. The final prediction is the average of all tree predictions.

$$\text{RandomForest}(x) = \frac{1}{N} \sum_{i=1}^{N} T_i(x) \tag{9}$$

Represents the averaging of predictions for regression tasks, where each tree T_i provides a numeric prediction.

DecisionTree: Decision Tree is a supervised learning algorithm that recursively splits the dataset based on features to make decisions. It predicts the target variable by traversing the tree from the root to a leaf node. Decision Trees are interpretable, handle both numerical and categorical data, and capture non-linear relationships in data. Here are some key aspects:

Let X_j be the selected feature at node j and threshold_j be the threshold for the decision rule. The decision rule can be expressed as:

$$\text{Node}_j \leq \text{threshold}_j \tag{10}$$

This indicates that if the value of the feature for a data point is less than or equal to the threshold, the left branch is taken; otherwise, the right branch is taken. The criterion for selecting the best feature and threshold for splitting can involve impurity measures. Let $\text{Impurity}(D)$ represent the impurity measure for a dataset D. The splitting criterion might be:

$$\text{SplitCriterion}(D) = \text{Impurity}(D) - \sum_{i=1}^{N} \left(\frac{|D_i|}{|D|} \cdot \text{Impurity}(D_i) \right) \tag{11}$$

where D_i represents the dataset in the i^{th} child node.

LightGBM is a gradient boosting framework that uses tree-based learning. It builds trees vertically, splitting the leaf with the maximum delta loss, leading to a more efficient and accurate model. LightGBM is known for its speed, efficiency, and ability to handle large datasets, making it suitable for distributed computing and parallel training.

The training objective in LightGBM can be described mathematically as minimizing the following objective function:

$$\text{Objective} = \sum_{i=1}^{N} \text{loss}(y_i, (\hat{y}_i)) + \sum_{k=1}^{K} \Omega(f_k), \tag{12}$$

where N is the number of samples, loss is a differentiable loss function, $(\hat{y}_i)$ is the predicted value, f_k represents the k^{th} tree, and $\Omega(f_k)$ is the regularization term for the k^{th} tree.

3.2.2.3.Deep learning Models

TabNet is an attention-based neural network designed for tabular data. It employs a combination of sequential and attention-based processing to capture complex relationships in tabular datasets. TabNet's advantages include interpretability, feature importance analysis, and effective handling of both categorical and numerical features in tabular data.

TabNet Structure: TabNet consists of multiple decision steps, each involving the following sub-steps.

Feature Selection: At each step, features are selectively chosen using attention mechanisms.

Attention Mechanism: Attention scores are calculated using the sparsemax activation, assigning importance to features based on their relevance to the task.

In each decision step t, the attention mask is computed using the sparsemax activation function:

$$a_{i,t} = \text{sparsemax}(f_{i,t}), \tag{13}$$

where $a_{i,t}$ is the attention score for feature i at step t, and $f_{i,t}$ is the logit before applying sparsemax. The sparsemax function is defined as:

$$\text{sparsemax}(z)_i = \max(z_i - \tau, 0), \tag{14}$$

where τ is a learnable parameter controlling the sparsity of the attention.

After the above steps, the model constructs a decision tree-like structure through a series of decision steps. Each step refines the features considered for making predictions.

Training Objective: The overall training objective includes a task-specific loss (e.g., cross-entropy for classification) and a sparsity penalty term:

$$\text{Objective} = \text{Task-specific Loss} + \lambda \cdot \text{Sparsity Penalty}, \tag{15}$$

where λ controls the strength of the sparsity penalty.

Sparsity Penalty: The sparsity penalty is applied to the attention mask to encourage sparsity:

$$\text{Sparsity Penalty} = \sum_{i,t} \|\text{sparsemax}(f_{i,t})\|_0, \tag{16}$$

promoting the model to focus on a limited set of features.

3.2.2.4.Feature Importance and Model's Performance evaluation

Based on trained models, including TabTransformer, FTTransformer, TabNet, XGBoost, AdaBoost, RandomForest, DecisionTree, and LightGBM, we analyze their impact on factors other than the CHD field in the CHD dataset by obtaining feature importance scores on the test dataset. Specifically, for the Transformer model, we use `linear_transformer.predict(test_dataset)['importances']` to retrieve feature importance scores (shown in Figure 4, $FIF_1, FIF_2, \ldots, FIF_{10}$), while for TabNet, XGBoost, AdaBoost, RandomForest, DecisionTree, and LightGBM, we use `feature_importances`. These models provide output values such as Gain, Cover, and Weight.

$$\text{Gain} = \frac{I}{N}, \tag{17}$$

where I represents the improvement in the loss function and N denotes the number of samples.

$$\text{Cover} = \frac{T}{N}, \tag{18}$$

where Tstands for the number of observations concerned by the feature.

$$\text{Weight} = \frac{S}{N}, \tag{19}$$

where S signifies the number of samples associated with the feature.
Subsequently, we compute the average scores of Gain, Cover, and Weight from each model on different features, retaining the top 10 features with the highest average scores.

$$D_{\text{new}} = D\left(F\left(T\left(A_1^i \sum(Gain, Cover, Weight)\right)\right)\right) \tag{20}$$

Where $A_1^i \sum(Gain, Cover, Weight)$ calculates the average feature scores for Gain, Cover, and Weight across all features, $T()$ is selecting the top 10 features for different models, $F()$ is determining the frequency of features Top 10 in different models, and $D()$ is obtaining a new dataset according to top features. By considering these consistently high-importance features, we use a frequency metric to identify the final Top features. After obtaining the Top features, we construct a new feature dataset and conduct a secondary modeling phase.

The model's performance is evaluated using a confusion matrix [41], accuracy, precision, recall, and F1-score [42].

Table 2: Confusion Matrix

		ACTUAL	
		Positive	**Negative**
PREDICTED	**Positive**	TP	FP
	Negative	FN	TN

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{21}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{22}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{23}$$

$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{24}$$

This step aims at seeking the best-performing model, preparing for the construction of an ensemble model.

3.2.2.5.Intergration Models

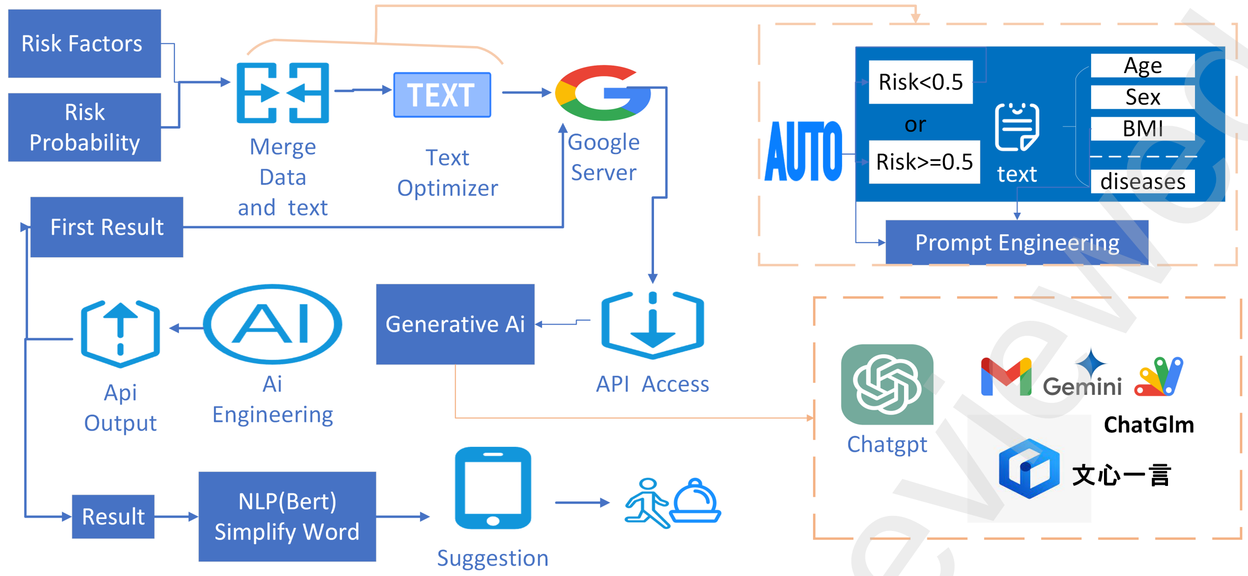
After obtaining the best models, for more precise results, this study will integrate models, as shown in Figure 4. Specific details are provided below. The ensemble of models will combine evaluations from the best machine learning models, including LightGBM, RandomForest, XGBoost, and TabNet, along with deep learning models, for risk probability prediction based on user-provided data. The script iterates through each model, extracting predicted probabilities,

$$P = \mathcal{A}_1^n(M_p), \tag{25}$$

where M_p is each model's probabilities (shown in Figure 4, $probability(1), \ldots, probability(n - i), \ldots, probability(n)$), and $\mathcal{A}$ calculates average scores for each model's probabilities. P is the average prediction. Feature importance and the most significant features are considered. The average prediction is computed, and if it surpasses the 0.5 threshold, the output indicates a high-risk probability in red; otherwise, it signifies a low-risk probability in green. Additionally, based on the input data, the script retrieves the scores for key factors from each model and identifies the top features obtained by each model.

$$F = \mathcal{T}_{\text{top}}\left(\mathcal{S}_1^n(M_1^n)_F\right), \tag{26}$$

where M_F is each model's features, S is sort feature values from high score to low score, F is the final feature values. The unique features are then deduplicated to obtain the latest personalized risk factors and their corresponding feature values. This comprehensive approach allows for a thorough risk assessment based on the contributions of multiple models and identifies the most influential features.



en

Figure 5: Large Language Model (LLM) Proposal Suggestion (Task 3)

Table 3: LLM Reply Result and Details

LLM Name	Run Time	Word Count	Important Word Count	Correlation Coefficient for Important Word	Whether Have Detail for Different Factors (Y/N)
Chatgpt	14.68	227	189	0.83	Y
Chatglm	15.44	215	168	0.78	N
Baidu	14.88	475	379	0.80	Y
Gemini Pro	12.51	314	243	0.62	N

3.3. Large Language Model (LLM) Proposal Suggestion(Task 3)

Firstly, Table 3 presents response time, word count, proportion of important words, and the provision of specific advice for different factors across four large language models (LLMs). Among them, ChatGPT exhibits a relatively longer runtime (14.68 seconds), with a response of 227 words containing 189 important words and a high correlation coefficient of 0.83. It also provides detailed advice for different factors. In contrast, ChatGLM has slightly inferior performance in terms of runtime (15.44 seconds), response word count (215), and important word count (168), and it does not offer detailed advice. Baidu's "Wenxin Yiyuan" shows a runtime of 14.88 seconds, a response of 475 words containing 379 important words, a high correlation coefficient of 0.80, and detailed advice for different factors. Gemini Pro has relatively lower runtime (12.51 seconds), response word count (314), and important word count (243), with a correlation coefficient of 0.62 and no detailed advice.

ChatGPT and Baidu stand out in responses with longer runtimes, higher word counts, and a significant proportion of important words. GPT3.5's superior performance in important word usage makes it suitable for the generative AI module in the advisory stage of the system.

As depicted in Figure 5, the user's process of obtaining optimal advice based on their data is detailed. Before interacting with the GPT3.5 API, risk probabilities, high-risk factors, high-risk values from Task 2, and all user inputs from Task 1 are prepared for integration. The prompt text is edited to transform it into a structured and coherent text format using Python's text combination functions (e.g., + operator). The interaction with GPT3.5 (initial feedback) involves passing the optimized text as input, initiating a conversation, and receiving initial feedback, typically comprising extensive text suggestions. To streamline and enhance the obtained suggestions, the initial results are simplified by modifying the prompt and transmitted again to GPT3.5 for more concise and intuitive feedback (Final Feedback). For optimal refinement, the BERT technique(details shown in below) in NLP is utilized to obtain and combine key text once again, resulting in the final and most refined outcome [43].

$$\text{L2\_norm}(\mathbf{v}) = \sqrt{\sum_{i=1}^{n} v_i^2} \tag{27}$$

Where: $\text{L2\_norm}(\mathbf{v})$ represents the $L2norm$ (Euclidean norm) of vector $\mathbf{v}$, where $\mathbf{v}$ denotes the embedding vector of a sentence, n is the dimension of the vector, v_i refers to the $i-th$ element of the vector $\mathbf{v}$, $\sum_{i=1}^{n}$ signifies the summation over elements from 1 to n, and $\sqrt{\ldots}$ denotes taking the square root.

4. Experiment

4.1. Dataset

We conducted this study using the dataset Behavioral Risk Factor Surveillance System (BRFSS)[1], which was initially obtained from the CDC. This dataset was derived from the CSV of the 2015 available dataset. The original dataset contains responses from 441,455 individuals with 22 features. These features were questions posed directly to participants or variables calculated from individual participant responses. This dataset contains 253,680 survey responses from the cleaned BRFSS 2015 and was used primarily for the binary classification of heart disease. It is displayed in Table 4.

Since there is a severe imbalance in the presence or absence of CHD in this dataset, where CHD is "no" for 229,787 individuals and CHD is "yes" for 23,893 individuals, the ratio reaches 10:1. Therefore, at this moment, the data with CHD 0 is selected based on the data situation with CHD 1. The specific number of rows is shown in Table 5.

Table 4: Dataset Detail

Dataset	**Label**	**Number**	**Total**
Heart Disease Health	CHD(Yes)	23,893	47,786
Indicators Dataset	CHD(No)	23,893	

4.2. result

This section presents the experimental results of the research methodology. The first presentation that is carried out is the presentation of the trained model for the comparison of the top10 of the feature values obtained from the evaluation and prediction of the dataset.

4.2.1. Feature Engineering

Table 6 and Figure 6 show the top 10 eigenvalues of feature importance obtained by applying all algorithms. Figure 6 shows that Age, High Blood Pressure, body mass index(BMI), income, high cholesterol, Health Degree(GenHlth), Physical Health, Sex(Gender), Mental Health, Stroke, Education, and diabetes appear very frequently in all models. Table 4 shows the top 10 essential features for predicting coronary heart disease (CHD) risk using machine learning models. Each row in the data corresponds to a specific feature, while each column represents the different models' assessment of the importance of that feature in CHD risk prediction.An important observation is the ranking of feature importance. The data demonstrate how each model ranks features based on their relevance to heart disease risk prediction. According to the model, the feature that ranks first in heart disease risk prediction is at the top and decreases in importance as the column moves down. Another aspect worth noting is the variability of the model. Different models often have different views of feature importance. For example, FTTransformer,XGBoost,TabNet ,consider GenHlth to be necessary, but in models such as RandomForest, DecisionTree, AdaBoost, GenHlth not have the same importance. In addition, the data revealed the concept of consistency. Some characteristics, such as Age, HighBP, and BMI were consistently important across multiple models. This consistency suggests these characteristics may be essential in predicting coronary heart disease risk. Unique insights can also be gained from these data. For example, RandomForest emphasizes education, while AdaBoost emphasizes smoking as a significant predictor. Ensemble models like RandomForest and AdaBoost integrate multiple decision trees or base models. Thus, given the consensus of many models, they provide a more comprehensive view of feature importance. To further validate the accuracy and robustness of these models, the following will continue model evaluation.The detail shown in Table 7.

Table 5: Dataset Features Description

id	Category	Describe	Value
1	CHD	Having previously indicated a history of Coronary Heart Disease (CHD) or myocardial infarction (MI) at any point.	1/0
2	HighBP	Adults who have been told they have high blood pressure by a doctor, nurse, or other health professional	1/0
3	BMI	Computed body mass index	Float[1-9999]
4	HighChol	Having High cholesterol or 0t	1/0
5	Smoker	Have you consumed a minimum of 100 cigarettes (5 packs) throughout your lifetime?	1/0
6	CholCheck	Having cholesterol Check or 0t	1/0
7	HvyAlcoholConsump	Heavy drinkers (adult men having more than 14 drinks per week and adult women having more than 7 drinks per week)	1/0
8	Stroke	(Ever told) (you had) a stroke.	1/0
9	PhysHlth	Consider your overall physical well-being, encompassing instances of physical ailments and injuries. Over the previous month, how many days did you experience subpar physical health?	Number of days [1-30]
10	MentHlth	Consider your psychological well-being, encompassing stress, depression, and emotional challenges. How many days out of the last 30 days did you experience poor mental health?	Number of days [1-30]
11	DiffWalk	Are you facing significant challenges when it comes to walking or ascending stairs?	1/0
12	Sex	Are you male or female?	Male/Female
13	Age	Real Age	Age
14	Fruits	Eating Fruit 1 or more times per day	1/0
15	Diabetic	Have you ever been informed about having diabetes? In case the answer is '1' and the person responding is female, inquire whether this was specifically during pregnancy. If the respondent mentions pre-diabetes or borderline diabetes, utilize response code 4.	2-No, But diabetes /0-No/1-diabetes
16	PhysActivity	In the previous month, apart from your usual work, did you engage in any physical pursuits or workouts like jogging, bodyweight exercises, golfing, tending to gardens, or purposeful walking?	1/0
17	GenHealth	Do you consider your overall health, on the whole, to be:	5-Excellent/4-Very good/3-Good/2-Fair/1-Poor
18	Education	What is the most advanced level of education you have finished?	Number of levels [1-6]
19	Income	Does your yearly household earnings encompass all origins of income? (In case the participant declines to provide income information at any level, mark it as "Refused.")	Number of degrees [1-10]
20	Veggies	Integrate vegetables into your diet at least once a day. Within the last year, did you experience a situation where you required medical attention but refrained due to financial constraints?	1/0
21	0DocbcCost	Integrate vegetables into your diet at least once a day. Within the last year, did you experience a situation where you required medical attention but refrained due to financial constraints?	1/0
22	AnyHealthcare	Do you possess any form of healthcare protection, encompassing health insurance, prepaid schemes like HMOs, or governmental programs such as Medicare or the Indian Health Service?	1/0

Table 6: Top 10 Features of Different Algorithms

Id	FTTransformer	TabNet	XGBoost	RandomForest	DecisionTree	AdaBoost	LightGBM
1	Sex	DiffWalk	HighBP	BMI	BMI	HighBP	Age
2	HighBP	Sex	DiffWalk	Age	Income	HighChol	BMI
3	MentHlth	HighBP	HighChol	GenHlth	Age	CholCheck	GenHlth
4	Age	HighChol	BMI	Income	PhysHlth	BMI	Income
5	HighChol	GenHlth	Stroke	HighBP	GenHlth	Smoker	PhysHlth
6	Education	Fruits	Sex	PhysHlth	Education	Stroke	MentHlth
7	Stroke	NoDocbcCost	Income	Education	MentHlth	Diabetes	Sex
8	PhysHlth	Diabetes	PhysHlth	MentHlth	Fruits	PhysActivity	Stroke
9	Smoker	HvyAlcoholholConsump	GenHlth	HighChol	PhysActivity	Fruits	Education
10	Income	Age	Age	DiffWalk	Diabetes	Veggies	Diabetes

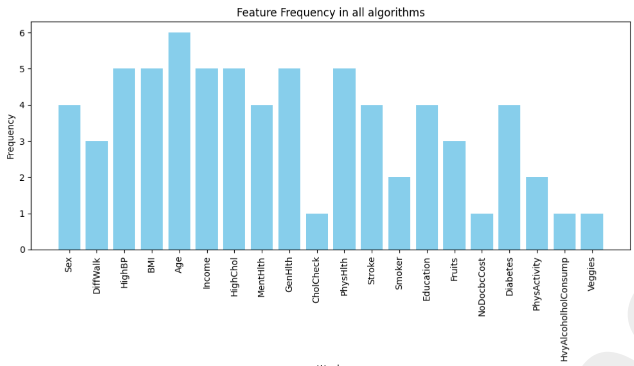


Figure 6: Top 10 eigenvalues of feature importance

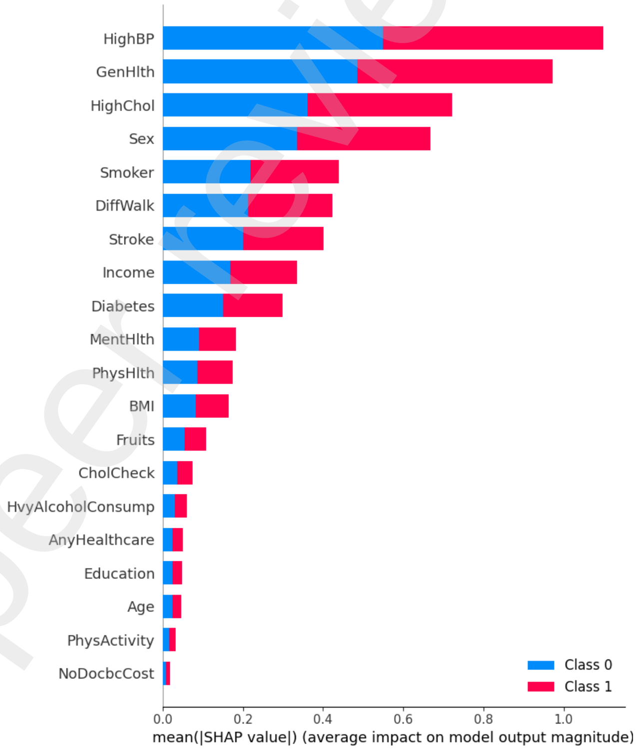


Figure 7: LightGBM

4.2.2. Performance Of Models

Based on the data in Table 7, we analyze the performance indexes of various algorithms. TabTransformer has an accuracy of 0.74, precision of 0.74, recall of 0.76, and F1-Score of 0.75. FTTransformer has an accuracy of 0.75, precision of 0.73, recall of 0.78, and F1-Score of 0.75. TabNet has the highest accuracy of 0.77. It has an accuracy of 0.75, recall of 0.81, and F1-Score of 0.78. XGBoost, AdaBoost, and LightGBM all have an accuracy of 0.77. They have similar precision, recall, and F1-Score values, with similar accuracy, precision, recall, and F1-Score. Score values are also similar, with a precision of about 0.76, a recall of about 0.80, and an F1-Score of about 0.78. RandomForest has an accuracy of 0.76, precision of 0.75, recall of 0.79, and F1-Score of 0.77. Its ROC-AUC score is 0.83. The DecisionTree algorithm has the lowest accuracy of 0.67, and its precision, recall, and F1-Score are all around 0.67. Based on the above analysis, XGBoost, AdaBoost, and LightGBM perform well in predicting CHD based on specific CHD risk factors. Model integration was performed based on these three models and was evaluated with a score of 0.77 for Accuracy, 0.76 for Precision, 0.81 for Recall, and 0.78 for F1-Score.

[1] https://www.kaggle.com/datasets/alexteboul/heart-disease-health-indicators-dataset

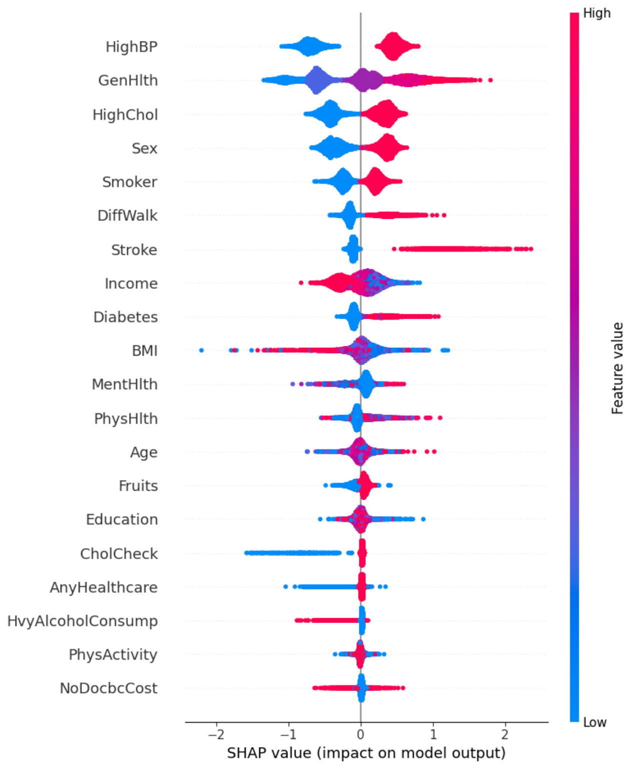


Figure 8: XGBoost

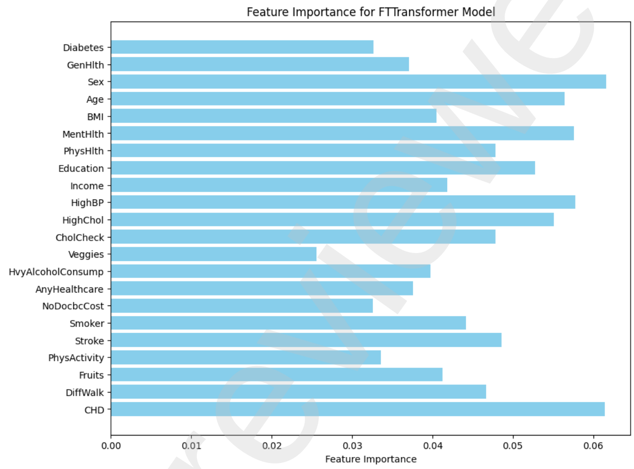


Figure 9: FTT

4.2.3. Large Language Model (LLM) Proposal Suggestion Result

Based on the data presented in Table 8, it is evident that the provided system (HCMS) demonstrated exceptional accuracy and was highly consistent with the person-centered philosophy. All participants received a NORISK(¡0.5) health prediction, demonstrating that the system was consistently able to accurately assess the user's level of health risk. The goal of the system is not only to provide individuals with information but also to empower them to make informed decisions about their health. By providing accurate predictions and refined guidance (Detail in Figure 10), it provides users with the knowledge and tools they need to take proactive steps to improve their health and lifestyle. Gathering user feedback, monitoring long-term health outcomes, and staying current with evolving research and medical guidelines are important steps in ensuring the continued success of a user-centered healthcare system. Figure 10 displays the results of the suggestion response (an example from 5 participants). Obtaining a probability of high risk (shown in the graph as 0.44), high-risk factors, and risk values (BMI 2, HighBp 0, $Genhealth_{e}xcellent$: 1), the system generates recommendations based on both the user's input and the predicted data, aligning with the expectations for suggestions across different indicators. Table 8 presents the system responsiveness and user outcome data for the five users. All

Table 7: Performance in different Algorithm Models

Algorithm	Accuracy	Precision	Recall	F1-Score
TabTransformer	0.74	0.74	0.76	0.75
FTTransformer	0.75	0.73	0.78	0.75
TabNet	0.77	0.75	0.81	0.78
XGBoost	0.77	0.76	0.80	0.78
AdaBoost	0.77	0.76	0.80	0.78
RandomForest	0.76	0.75	0.79	0.77
DecisionTree	0.67	0.68	0.67	0.67
LightGBM	0.77	0.76	0.81	0.78
Modeling Integration	0.77	0.76	0.81	0.78

Table 8: System Response Result.(Notes: U-User, H-Height, W-weight(Kg), A-Age, G-Gender, S-Stroke, PA-Physical Activity, C-Cholesterol(mg/dl), BP-Blood Pressure(systolic): (mm Hg), E-Education, HD-Health Degree, MH-Mental Health(1-30 days), PH-Physical Health(1-30 days), D-Diabetes, I-Income($ per year), No-No Risk, F-Female, M-Male,AVS-Average Response Speed)

U	H	W	A	G	S	PA	C	BP	E	HD	MH	PH	D	I	Result	Suggestion	AVS
U1	1.8	80	25	M	NO	YES	120	110	Degree	Good	30	30	NO	80000	0.43	Tomatoes, Walking	0.36s
U2	1.62	60	24	F	NO	YES	114	105	Master	Good	30	30	NO	100000	0.35	Salad, Quiet Place	0.34s
U3	1.66	64	24	F	NO	YES	113	109	Master	Good	30	30	NO	85000	0.28	Jogging, Banana	0.32s
U4	1.78	68	23	F	NO	YES	108	107	Master	Good	30	30	NO	67000	0.37	Swimming, Walk	0.35s
U5	1.77	74	24	F	NO	YES	104	112	Master	Good	30	30	NO	43000	0.45	Fruit, Fresh Air	0.34s

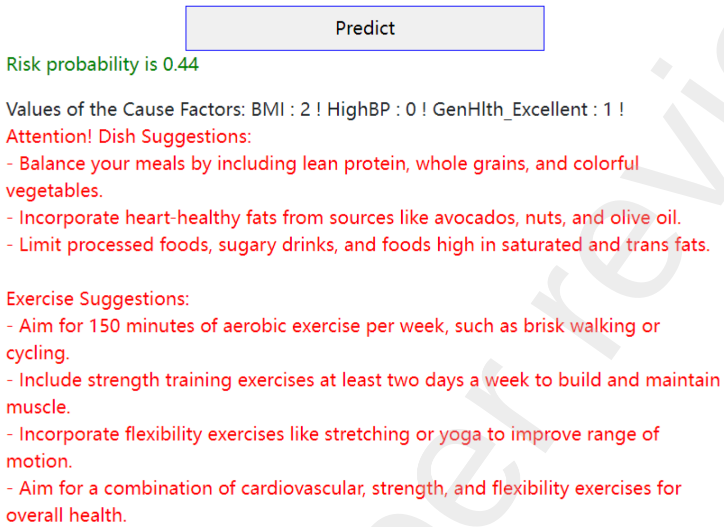


Figure 10: Example of Getting Advice

users, namely User 1, User 2, User 3, User 4, and User 5, received a risk result of ¡0.5 and corresponding suggestions, indicating positive outcomes. The average response time varied slightly among users, ranging from 0.32 to 0.36 seconds. These results suggest that the system generally responds quickly, although there may be subtle differences in response times.

5. Conclusion And Future Work

In this article, we propose a human-centered intelligent monitoring framework that utilizes machine learning to assess risk factors and enhance the accuracy of heart disease prediction. The framework incorporates GPT3.5, a widely-used AI technology, to provide precise advice for quick and accurate diagnosis by doctors and to assist patients in specifying a sensible diet aligned with healthy living and exercise standards. Various key aspects are discussed, including obtaining input data from the population, storing datasets with data processing techniques, defining high-risk factors through machine learning and transformer modeling, risk prediction using optimal modeling, and recommending risk outcomes with the GPT3.5 technique.

After obtaining features such as Age, High Blood Pressure, BMI, income, high cholesterol, Health Degree (GenHlth), Physical Health, Gender, Mental Health, Stroke, Education, and diabetes, it was determined that these are the most critical risk factors. XGBoost, AdaBoost, and LightGBM emerged as the best-performing algorithms among all models. The proposed framework integrates these algorithms for a comprehensive CHD risk assessment, yielding the desired results.

However, due to equipment limitations, experiments with voice input, direct prediction of other diseases, and utilization of image data were not conducted precisely. Future work will explore the use of NLP, LLM, and deep learning for preliminary risk predictions based on images. Additionally, video reconnaissance of user dynamics and speech recognition techniques will be applied to enhance data extraction.

References

[1] P. Deng, Y. Fu, M. Chen, D. Wang, L. Si, Temporal trends in inequalities of the burden of cardiovascular disease across 186 countries and territories, International Journal for Equity in Health 22 (1) (2023) 164.

[2] D. Prabhakaran, P. Jeemon, A. Roy, Cardiovascular diseases in india: current epidemiology and future directions, Circulation 133 (16) (2016) 1605–1620.

[3] J. Li, X. Lu, Optimising targeted dietary intervention strategies based on regional disparities of the related burden of cardiovascular disease in china, The Lancet Regional Health–Western Pacific (2023).

[4] C. Wuche, The cardiovascular system and associated disorders, British Journal of Nursing 31 (17) (2022) 886–892.

[5] R. K. Upadhyay, High cholesterol disorders, myocardial infarction and its therapeutics, World Journal of Cardiovascular Diseases 13 (8) (2023) 433–469.

[6] T. Ramesh, U. K. Lilhore, M. Poongodi, S. Simaiya, A. Kaur, M. Hamdi, Predictive analysis of heart diseases with machine learning approaches, Malaysian Journal of Computer Science (2022) 132–148.

[7] G. Ramkumar, J. Seetha, R. Priyadarshini, M. Gopila, G. Saranya, Iot-based patient monitoring system for predicting heart disease using deep learning, Measurement (2023) 113235.

[8] S. Byiringiro, N. Nyirimanzi, J. Mucumbitsi, E. R. Kamanzi, J. Swain, Cardiac surgery: increasing access in low-and middle-income countries, Current Cardiology Reports 22 (2020) 1–7.

[9] H. Ahmed, E. M. Younis, A. Hendawi, A. A. Ali, Heart disease identification from patients' social posts, machine learning solution on spark, Future Generation Computer Systems 111 (2020) 714–722.

[10] P. Melin, I. Miramontes, G. Prado-Arechiga, A hybrid model based on modular neural networks and fuzzy systems for classification of blood pressure and hypertension risk diagnosis, Expert Systems with Applications 107 (2018) 146–164.

[11] S. Nazari, M. Fallah, H. Kazemipoor, A. Salehipour, A fuzzy inference-fuzzy analytic hierarchy process-based clinical decision support system for diagnosis of heart diseases, Expert Systems with Applications 95 (2018) 261–271.

[12] J. Yan, J. Tian, H. Yang, G. Han, Y. Liu, H. He, Q. Han, Y. Zhang, A clinical decision support system for predicting coronary artery stenosis in patients with suspected coronary heart disease, Computers in Biology and Medicine 151 (2022) 106300.

[13] F. Ali, S. El-Sappagh, S. R. Islam, D. Kwak, A. Ali, M. Imran, K.-S. Kwak, A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion, Information Fusion 63 (2020) 208–222.

[14] C. Ardito, P. Buono, D. Caivano, M. F. Costabile, R. Lanzilotti, Y. Dittrich, Human-centered design in industry: lessons from the trenches, Computer 47 (12) (2014) 86–89.

[15] U. R. Acharya, K. S. Vidya, D. N. Ghista, W. J. E. Lim, F. Molinari, M. Sankaranarayanan, Computer-aided diagnosis of diabetic subjects by heart rate variability signals using discrete wavelet transform method, Knowledge-based systems 81 (2015) 56–64.

[16] A. R. N. Suneetha, T. Mahalngam, Fine tuning bert based approach for cardiovascular disease diagnosis, International Journal of Intelligent Systems and Applications in Engineering 11 (6s) (2023) 59–66.

[17] S. Meda, R. Bhogapathi, Identification of heart disease using fuzzy neural genetic algorithm with data mining techniques, Adv Model Anal B 61 (2018) 99–105.

[18] W. Wan Zunaidi, R. Saedudin, Z. Ali Shah, S. Kasim, C. Sen Seah, M. Abdurohman, Performances analysis of heart disease dataset using different data mining classifications, Int J Adv Sci Eng Inf Technol 8 (2018) 2677.

[19] R. Fadnavis, K. Dhore, D. Gupta, J. Waghmare, D. Kosankar, Heart disease prediction using data mining, J Phys Conf Ser 1913 (2021) 012099.

[20] R. K. AL-Taie, B. Saleh, A. Falih Saedi, L. Salman, Analysis of weka data mining algorithms bayes net, random forest, mlp and smo for heart disease prediction system: A case study in iraq, Int J Electr Comput Eng IJECE 11 (2021) 5229–5239.

[21] G. Wang, L. Luo, X. Zhao, Diagnosis of arrhythmia for patients with occult coronary heart disease guided by intracavitary electrocardiogram under data mining algorithm, J Healthc Eng 2021 (2021) 1640870.

[22] A. Gonsalves, F. Thabtah, R. Mohammad, G. Singh, Prediction of coronary heart disease using machine learning: An experimental analysis, Proc. 2019 3rd Int. Conf. Deep Learn. Technol. (2019) 51–6.

[23] E. Daoud, Comparison between xgboost, lightgbm and catboost using a home credit dataset 13 (2019).

[24] Z. Du, Y. Yang, J. Zheng, Q. Li, D. Lin, Y. Li, Accurate prediction of coronary heart disease for patients with hypertension from electronic health records with big data and machine-learning methods: Model development and performance evaluation, JMIR Med Inform 8 (2020) e17257.

[25] R. Nakanishi, P. Slomka, R. Rios, J. Betancur, M. Blaha, K. Nasir, Machine learning adds to clinical and cac assessments in predicting 10-year chd and cvd deaths, JACC Cardiovasc Imaging 14 (2021) 615–25.

[26] T. Shen, D. Liu, Z. Lin, C. Ren, W. Zhao, W. Gao, A machine learning model to predict cardiovascular events during exercise evaluation in patients with coronaryheart disease (2022).

[27] N. Absar, E. Das, S. Shoma, M. Khandaker, M. Miraz, M. Faruque, The efficacy of machine-learning-supported smart system for heart disease prediction, Healthcare 10 (2022) 1137.

[28] J. Hassan ChAU, Iqbal, R. Irfan, S. Hussain, A. Algarni, S. Bukhari, Effectively predicting the presence of coronary heart disease using machine learning classifiers, Sensors 22 (2022) 7227.

[29] S. Zhang, Y. Yuan, Z. Yao, J. Yang, X. Wang, J. Tian, Coronary artery disease detection model based on class balancing methods and lightgbm algorithm, Electronics 11 (2022) 1495.

[30] K. Wankhede, B. Wukkadada, S. Rajesh, S. Nair, Machine learning techniques for heart disease prediction (2023) 28–33.

[31] R. Yilmaz, F. Yağin, Early detection of coronary heart disease based on machine learning methods, Med Rec 4 (2022) 1–6.

[32] M. Hassanuzzaman, N. Hasan, M. Mamun, M. Alkhodari, K. Ahmed, e. a. Khandoker, AH, Recognition of pediatric congenital heart diseases by using phonocardiogram signals and transformer-based neural networks, 2023 45th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBC (2023) 1–4doi:10.1109/EMBC40787.2023.10340370.

[33] M. Gangadhar, K. Sai, S. Kumar, K. Kumar, M. Kavitha, S. Aravinth, Machine learning and deep learning techniques on accurate risk prediction of coronary heart disease, 2023 7th Int. Conf. Comput. Methodol. Commun. ICCMC (2023) 227–32doi:10.1109/ICCMC56507.2023.10083756.

[34] H. Yun, J. Lim, E. Lee, Genetic risk score for prediction of coronary heart disease in the korean genome and epidemiology study, Rev Cardiovasc Med 24 (2023) 102. doi:10.31083/j.rcm2404102.

[35] K. Wang, H. Meng, X. Wang, Application of vision-series transformer in screening for coronary heart diseases using coronary ct angiography (2023) 421–5doi:10.1145/3603781.3603858.

[36] G. Kim, M. Yang, G.-H. Kim, S.-H. Eom, T.-S. Lee, S. Park, Predicting heart failure prognosis using deep learning based on ft-transformer (2023) 932–6doi:10.1109/ICUFN57995.2023.10200998.

[37] K. Budholiya, S. Shrivastava, V. Sharma, An optimized xgboost based diagnostic system for effective prediction of heart disease, J King Saud Univ - Comput Inf Sci 34 (2022) 4514–23. doi:10.1016/j.jksuci.2020.10.013.

[38] Y. Chen, H. Li, H. Dou, H. Wen, Y. Dong, Prediction and visual analysis of food safety risk based on tabnet-gra, Foods 12 (2023) 3113. doi:10.3390/foods12163113.

[39] L. Maison, Y. Estève, Some voices are too common: Building fair speech recognition systems using the common voice dataset (2023).

[40] S. Uddin, A. Khan, M. Hossain, M. Moni, Comparing different supervised machine learning algorithms for disease prediction, BMC Med Inform Decis Mak 19 (2019) 281. doi:10.1186/s12911-019-1004-8.

[41] M. Heydarian, T. Doyle, R. Samavi, Mlcm: Multi-label confusion matrix, IEEE Access 10 (2022) 19083–95. doi:10.1109/ACCESS.2022.3151048.

[42] T. Fawcett, An introduction to roc analysis, Pattern Recognit Lett 27 (2006) 861–74. doi:10.1016/j.patrec.2005.10.010.

[43] A. Bulat, G. Tzimiropoulos, Lasp: Text-to-text optimization for language-aware soft prompting of vision & language models (2023) 23232–41doi:10.1109/CVPR52729.2023.02225.