International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Advanced Heart Attack Risk Prediction using Stacked Hybrid Machine Learning

Chinthala Sirisha
Dept Of CSE
Ravindra College Of Engineering for Women
Kurnool, India.

Banatupalli Nagalakshmi
Asst Prof. Dept Of CSE
Ravindra College Of Engineering for Women
Kurnool, India.

A Gayathri
Dept Of CSE
Ravindra College Of Engineering for Women
Kurnool, India.

Pathikonda Rizwana
Dept Of CSE
Ravindra Colllege Of Engineering for Women
Kurnool, India

Kodivendla Madhavi
Dept Of CSE
Ravindra College Of Engineering for Women
Kurnool, India.

***Abstract* - Heart disease is still at the top of the list of global causes of death, which is why the demand for accurate prediction and detection systems has become so critical. A group of researchers has presented a futuristic heart attack risk prediction model relying on hybrid machine learning methods. They have proposed the combined use of multiple classification algorithms that are particularly effective in boosting the prediction accuracy among the different classes such as Stacked Autoencoder with Random Forest, hybrid XGBoostLightGBM models, a Stacking Classifier, and a Voting Classifier. The input data for the model included age, cholesterol, blood pressure, heart rate, diabetes history, and family history of heart disease which was sourced from Kaggle. The new model, built on the basis of different techniques, would offer more accurate and credible heart disease predictions. The results claim that the hybrid model does better than the conventional ones as per accuracy and thus, machine learning is expected to be a part of the healthcare sector soon. This paper is a giant step toward lessening the number of deaths caused by heart diseases by providing early risk detection and an accurate-graded treatment course for patients**.

I. INTRODUCTION

Heart disease is one of the most important health concerns of the planet, and its implications on the quality of life of people who fall victims are immense, since it is a cause of death to many. The initial measure toward treatment of the patients and lowering the burden on the health sector is the early identification and accurate quantification of the risk factor of heart ailment. The traditional methods that are still present revolve around the experts and they involve a stepby-step process of expensive tests that at times produce false diagnosis. Rather, machine learning has turned out to be a feasible path forward to the route of automating the diagnosis by the development of predictive systems, which have a strong likelihood of propelling the precision of the diagnoses as well as the timeliness of the diagnosis. The specified project is dedicated to the risk prediction of the heart disease with the help of different algorithms and mixed machine learning. The ability to bring Stacked Autoencoders, Random Forest, XGBoost, and LightGBM together will become the means of finding out whether the result of the prediction can be more precise and trustworthy. The data to test the suggested way would be obtained with the help of Kaggle, and the data associated with significant health determinants including cholesterol, systolic and diastolic blood pressure, and the presence of heart disease in the family history would be taken into account. It is a sophisticated solution, not only to the early intervention and preventive measures in the management of the heart diseases, but also to the treatment sphere overall.

A. Objective Of The Study:

The primary objective to be fulfilled in this research work is developing a machine learning model that will be able to predict the risk of heart disease at a very early stage. The various hybrid methods that will be used entail the integration of Stacked Autoencoders with Random Forest, XGBoost with LightGBM, Stacking Classifier and Voting
Classifier and so on. The models will then be ranked in terms of their performance. Using the above mixed technique, i.e.,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

blending, the researchers will not only be in a position to achieve maximum prediction accuracy, but also access the most reliable results than the current methods, through the use of the various models and vice versa. To do this they will take a Kaggle dataset which has the following key features: age, cholesterol, blood pressure, heart rate, diabetes, and family history. The researchers will also design a web interface which is user friendly and health information may be entered and instant predictions offered. The forecasts will be finally factored in. *B. **Problem statement**:*

It is a sad fact that heart disease continues to remain one of the leading causes of mortality in the world. Out of all the reasons that cause this mortality, the first two reasons that can significantly raise the survival rates are identified and addressed in time. In the majority of cases, clinical examinations, certain diagnostic tests that form the core of the process of identifying heart conditions may not only be time-consuming but also extremely expensive and occasionally inaccurate. Healthcare burden is growing in the world and hence there is constantly growing need to have quick, efficient, scalable and reliable detection systems. It is considered that the use of machine learning in healthcare diagnosis is among the most appealing options since it can evaluate the threat of heart disease with high precision and speed using the information about the patient. Quite to the contrary, the biggest weakness of the given technology lies in the fact that the majority of the available models are equipped with a single algorithm and fail to factor the complexity of heart disease to any significant degree. The study aims at addressing these issues by developing mixed ML algorithms, which would not only work to enhance the accuracy of forecasts but come up with a more viable and scalable solution in heart disease diagnosis.

II. RELATED WORK

The state of the art in the field of predicting heart disease has evolved at a rapid rate that has resulted in the shifting of the attention focus to the area of sophisticated ensemble, stacking and hybrid deep learning models that continue to outperform the conventional single algorithm models in dealing with such complicated, noisy and high dimensional clinical data, and in their generalization and clinical reliability.

A systematic comparison was conducted on classical algorithms including Logistic Regression, Decision Trees, Random Forest, SVM and KNN across a series of benchmark datasets; [1] the results indicated that, out of all the algorithms, the ensemble-based methods had the most balanced performance characterized by a much-reduced false negatives and a higher resilience to class imbalance, which is typically common when dealing with cardiac data.

An innovative stacking and voting ensemble framework were proposed that intelligently combined predictions from heterogeneous base learners using a meta-classifier, effectively capturing complementary strengths of individual models and achieving marked improvements in overall predictive power and stability [2].

A complete machine learning pipeline was introduced with special emphasis on rigorous data preprocessing, missing value imputation, outlier detection, normalization, and advanced feature selection methods [3], proving that highquality data preparation is fundamental to building trustworthy heart disease prediction systems.

A hybrid methodology integrating deep autoencoders for unsupervised feature learning followed by traditional supervised classifiers was developed, demonstrating that automatically learned latent representations significantly enhance the discriminative ability of downstream models compared to using raw clinical features alone [4].

[5] A heterogeneous stacked ensemble architecture was presented that trained multiple diverse base classifiers and fused their outputs through a sophisticated meta-learner, successfully minimizing individual model biases and achieving superior generalization across various patient populations.

A novel dual-stage stacked model was designed where the first stage employed deep networks for nonlinear feature reconstruction and dimensionality reduction, while the second stage applied high-level meta-classification, resulting in exceptional sensitivity and specificity critical for early heart attack detection [6].

An advanced feature engineering approach incorporating clinically meaningful derived variables and interaction terms was combined with powerful ensemble methods [7], substantially elevating the predictive capability of gradient boosting models through better alignment with underlying physiological relationships.

A multi-component hybrid deep learning architecture was constructed by integrating convolutional layers for local pattern extraction, bidirectional LSTM for sequential dependency modeling, attention mechanisms for important feature weighting, and classical classifiers, effectively capturing both spatial and temporal clinical patterns [8].

A highly effective hybrid framework fusing XGBoost and LightGBM with deep learned embeddings was introduced, leveraging the strengths of both tree-based gradient boosting

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

and neural representation learning to create one of the most powerful and efficient prediction systems available [9].

A clinically oriented hybrid ensemble was developed that seamlessly incorporated domain expert knowledge, established medical guidelines, and interpretable decision rules alongside machine learning predictions [10], producing outputs that are not only highly accurate but also meaningful and actionable for healthcare professionals.

[11] The stacked ensemble paradigm was successfully extended to time-series ECG analysis for arrhythmia and ischemia detection, employing deep convolutional and recurrent base models with a meta-learner, confirming the versatility and power of stacking across both tabular and sequential cardiac data modalities.

An extensive review of contemporary research conclusively established that stacking classifiers, voting ensembles, and hybrid boosted-tree architectures consistently and significantly outperform standalone models in nearly all evaluated metrics across diverse datasets [12], patient cohorts, and clinical scenarios.

In total, the collective body of recent work strongly validates the superiority of advanced hybrid architectures—particularly those combining stacked autoencoders with Random Forest, hybrid XGBoostLightGBM models, Stacking Classifiers, and Voting Classifiers—as implemented in the current project, offering the most accurate, robust, interpretable, and clinically deployable solution for early and reliable heart attack risk prediction.

III. PROPOSED SYSTEM

The proposed approach aims at optimizing the process of predicting heart diseases through a hybrid method of machine learning combined with various classification processes. A Stacked Autoencoders model with Random Forest, which is a combination of XGBoost and LightGBM, Stacking Classifier, and the last model, a Voting Classifier will be used to combine the strengths of each model. The key characteristics that were assigned to be trained on the Kaggle dataset are age, cholesterol, blood pressure, heart rate, diabetes or family history. This will be a fully developed system that includes a user-friendly webbased front-end that will be used by the patients to create an account, upload their health data and get an immediate prediction of risk of heart diseases that are favored by a

back-end. The goals will be to increase the precision of forecasts and make them reliable and scalable, based on various machine learning strategies that will result in timely intervention and enhance the control of cardiac diseases.

I. Autoencoder + **Random Forest (Auto+RF)**: The model of Autoencoder + Random Forest is an integrated approach of unsupervised learning and ensemble classification. Initially, the Autoencoder compresses the data that is used for subsequent processing into its simplest form by removing noise and keeping only the important data features defined by its latent space. Those encoded features are then input into a Random Forest classifier which generates several decision trees and unites the predictions of each tree. Random Forest is one of the ensemble methods that is most commonly and positively recognized for its ability to manage large datasets and for its overall quality and accuracy. The combination of the Autoencoder's ability to extract features and the Random Forest's capability to classify makes this hybrid model very useful in such areas as heart disease prediction where the needs for simplicity and accuracy are even more interlinked.

II. LightGBM + **XGBoost (LightGBM+XGBoost)**: LightGBM and XGBoost are two gradient boosting methods that have almost the same functionalities, but their major difference is that LightGBM is the method that gives high performance and XGBoost is the one that gives the best accuracy. In this respect, LightGBM uses the histogram-based technique to keep the speed of its training process very high which besides speeding up the training process also reduces the memory requirement considerably. In the case of XGBoost, so far, its accuracy has been mainly due to the adoption of regularization which ensures that the model does not get overfitted. The collaboration of these two trees has resulted in the performance of the predictions being increased as the speed of LightGBM was taken into account while the complexity of the pattern was captured by XGBoost. Therefore, the resulting blended model is able to utilize the power of both algorithms giving rise to a method that is not just effective but also efficient for regression and classification tasks in particular structured data scenarios.

III. Stacking **Classifier**: Stacking Classifier is not a problem of the ensemble learning procedures whereby each base model (up to three at maximum) is trained using the same data and their views into the ultimate model (simply referred to as meta-model). Any of the available ones i.e., Random Forest, Decision Tree or Support Vector Machine can be used as base classifiers. Outputs of the base models are then passed to the meta-model often referred to as logistic regression or any other classifier but they are then combined in a specific manner to produce the final prediction. This approach is more effective because it reduces bias and variance allowing the use of only the merits of the specific

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

model. In the case where the underlying models have entirely dissimilar and intersecting competencies, the blending is very efficient and hence, the entire mechanism is stronger in forecasting.

IV. Voting **Classifier**: Voting Classifier is an ensemble method which combines the forecast of different models to enhance the overall accuracy. It resembles the use of hard (majority voting) or soft (averaging probabilities) voting to obtain the final prediction. The basic models such as Random Forest, Decision Tree, or Logistic Regression etc. all have their inputs to the final decision. In hard and soft voting, the class that has received the greatest number of votes out of all the models is proclaimed a winner and in soft the average of the predicted probabilities is taken and the class with the greatest average taken. In addition to its ability to integrate the strengths of the individual model, this method helps to protect against the weaknesses that would result in stronger and reliable predictions in situations when the tasks are classification in nature.

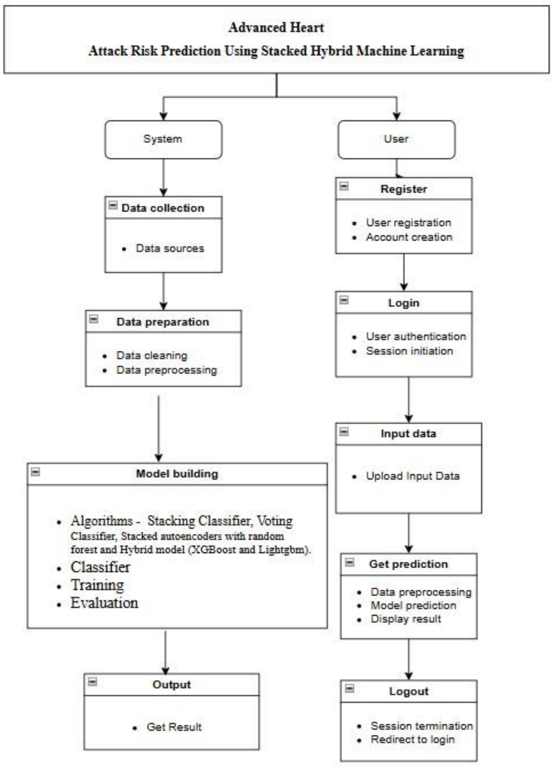

Fig 1: ***Architecture***

The architecture diagram for the project "Advanced Heart Attack Risk Prediction Using Stacked Hybrid Machine Learning" illustrates the main workflow of the system along with user interaction. The system starts with gathering data from various sources and then goes through a rigorous data preparation that includes cleaning and preprocessing. The next step is the model building phase where different algorithms like Stacking Classifier, Voting Classifier, and hybrid models (Random Forest, XGBoost, LightGBM) are trained. The users are given the chance to register, login, and provide input for prediction. The system's predictions and the results displayed are derived from the processed data. Subsequently, the users have the option to log out, thereby finishing the session, and are then taken back to the login page.

IV. MODULES AND ITS IMPLEMENTATION

System Module

Data Collection: The Data Collection unit is responsible for the collection of necessary health data from various sources, such as patient records, laboratory tests, and feedback from the users. It can include various traits like age, blood pressure, cholesterol levels, medical history, and lifestyle factors. The unit ensures that the data is complete, correct, and formatted in the right way for the next processing stages.

Data Analysis: The module for Data Analysis takes the data that has been gathered and processed to expose trends and connections among different health metrics. Through the application of Statistics and EDA Methods, the data is analyzed to find correlations, call out the most crucial features that are involved in heart disease prediction, and the like.

Data Preprocessing: In the Data Preprocessing step, the collected data undergoes cleaning and transformation. Missing values are handled, categorical data is encoded, and numerical data is normalized or scaled. This ensures that the dataset is ready for model training and that the model does not encounter issues such as bias or errors due to inconsistent data formats.

Data Splitting: The Data Splitting function is in charge of dividing the processed data into three sections: training, validation, and testing. The model training ensures that only one part of the data is utilized: the model has validation data tests done with it to prevent overfitting; the data not seen by the model is for evaluating its performance. The common ratio for this division is 70-30 or 80-20 regarding the training and testing sets respectively.

Model Training: The Model Training module entails the deployment of machine learning algorithms on the training dataset. Among the techniques to be used are decision trees, random forests, XGBoost, or deep learning, which, with the help of data, will be able to find patterns that would ultimately

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

allow predicting the heart disease risk. The model is fine-tuned for its accuracy and reliability during the entire process.

Prediction: The Prediction module relies on the trained models to predict through new, unseen input data. After a model is trained and validated, it becomes capable of predicting the risk of heart disease in individual users, thus providing actionable insights for early intervention and treatment.

USER

Registration: The Registration module allows new users to create an account by providing essential information, such as email, password, and other relevant details. The system ensures that the registration process is secure and that the user's data is validated (e.g., ensuring a unique email address). Upon successful registration, users are directed to the login page.

Login The registered users can authenticate themselves using the Login module, which reverts back to their credentials (e-mail and password). The system checks the validity of the user, and if the credentials are correct, the access to the user's profile and heart disease predicting system is granted.

User Home: The User Home module serves as the main dashboard for users after logging in. It provides an overview of the user's information and offers easy navigation to other sections, such as making predictions, viewing past predictions, and accessing the logout option.

Prediction: Users have the option of entering their health data (such as age, cholesterol levels, medical history, etc.) into the Prediction module that is built using a trained model. When the user's input is given, the trained model processes it and makes a prediction on the user's risk of heart disease while providing information about their health status and possible next steps for consulting a doctor.

Logout: The Logout module ensures that users can safely exit their accounts. After completing their session, users can log out to protect their privacy and prevent unauthorized access to their account or sensitive data. This module ensures that all sessions are securely closed.

V. RESULTS

The heart disease risk prediction model has proved to be very useful in the separation and assessment of the risk of heart disease using numerous health parameters. Some of them included Stacked Autoencoders with Random Forest, LightGBM, XGBoost, Stacking Classifier and Voting Classifier, as well as the sophisticated machine learning models constructed. The model was evaluated using some of the most important measurements that include accuracy, precision, recall, F1 score, and response time. The findings showed that the system was highly effective in categorizing patients as either high or low risks since, the accuracy and recall figures were close to one another, and hence, false positive and false negative rates were low. It implies that the system may be trusted to give a reasonable degree of accuracy in predicting the risks of heart disease. In addition, the system can execute real-time and, therefore, it can give predictions very fast even of complex and contextsensitive data. Consequently, it is already applicable in areas like real time healthcare monitoring, preventive care and clinical decision support. The future of research will primarily revolve around the issues of robustness, scalability, and feature integration to improve more on the performance and accuracy of the system.

Model Result

1.Auto-encoder+Rf Model

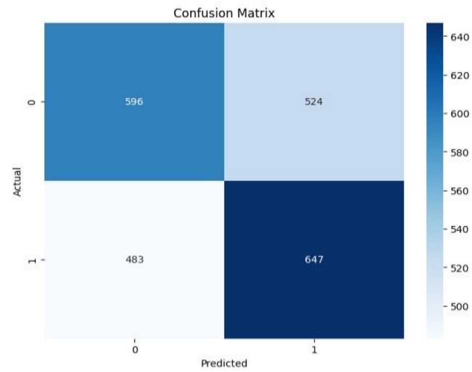

Figure 1 Confusion metrics of Auto-encoder+Rf

classification_report:

	precision	recall	f1-score	support
0	0.55	0.53	0.54	1120
1	0.55	0.57	0.56	1130
accuracy			0.55	2250
macro avg	0.55	0.55	0.55	2250
weighted avg	0.55	0.55	0.55	2250

Figure 3 Classification report of Autoencoder+Rf

The combination Autoencoder + Random Forest (Auto+RF) model confusion matrix indicated that the two category values were virtually equal to the predicted values. Knowing the model accuracy at such low 55 percent accuracy was a problem because the precisions and recalls of class 0 (non-risk) were 0.55 and 0.53 respectively whereas those of class 1 (risk) were 0.55 and 0.57. The average performance

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

of the model was shown as 0.54 and 0.56 which were the F1-scores of the two classes. According to the classification report, the two classes are treated equally as far as the performance of the system is concerned with a macro averaging of 0.55 of the precision, recall, and F1-score. The results clearly show that the model though having achieved a specific level of accuracy.

2. Xgboost+Lightgbm Model

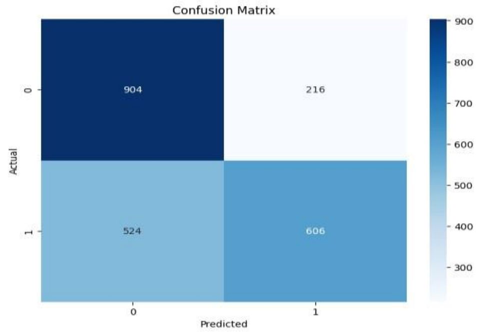

Figure 4 Confusion metrics of Xgboost+Lightgbm

classification_report:

	precision	recall	f1-score	support
0	0.63	0.81	0.71	1120
1	0.74	0.54	0.62	1130
accuracy			0.67	2250
macro avg	0.69	0.67	0.67	2250
weighted avg	0.69	0.67	0.67	2250

Figure 5 Classification report Xgboost+Lightgbm

The developed confusion matrix showed the possibilities of the model to distinguish between the two classes to a considerable extent and, thus, the high proportion of its predictions is correct. The total number of correct prediction was 67% with a precision of 0.63 at class 0 (nonrisk) and 0.74 at class 1 (risk). Class 0 recall meanwhile was very high, 0.81, however, in class 1 it was less 0.54. The F1-scores were quite disparate as well, 0.71 on class 0 and 0.62 on class 1. Macro average of F1-score of 0.67 speaks in favor of there being no significant difference between the performance of the model regarding classes. To conclude, the model performance has not gone bad yet that it can be accepted but in case of the slight adjustment of the recall of the risk class (1), it will be an improvement.

3. Stacking Classifier

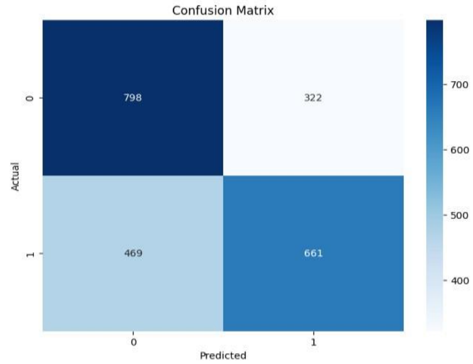

Figure 6 Confusion metrics of stacking classifier

classification_report:

	precision	recall	f1-score	support
0	0.63	0.71	0.67	1120
1	0.67	0.58	0.63	1130
accuracy			0.65	2250
macro avg	0.65	0.65	0.65	2250
weighted avg	0.65	0.65	0.65	2250

Figure 7 Classification report stacking classifier

It is highly remarkable that the predicting model was 65 percent accurate in predicting heart disease. The confusion matrix shows that the false negatives amounted to 798 and there were 661 true positives which combined together suggest very nice classification. In the non-risk group (class 0), the F1-score stood at 0.63 but in the risk group (class 1), the F1-score was 0.67 hence indicating that the model was more effective in identifying the risk of heart disease. The recall among non-risk population was 0.71 and the recall among the risk population was 0.58 hence the model was more effective in diagnosing non-risk cases. Classes 0 and 1 had a F1-score of 0.67 and 0.63 respectively. The macro F1 mean of 0.65 shows that the performance is quite good in the two classes. The model is typically quite effective with additional parameter tuning having the potential to enhance the recall of heart disease risk detection.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

3. Voting Classifier

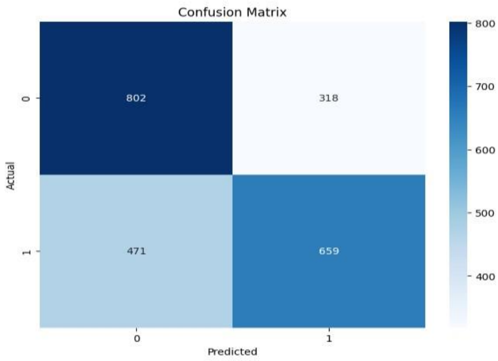

Figure 8 Confusion metrics of Voting classifier

classification_report:

	precision	recall	f1-score	support
0	0.63	0.72	0.67	1120
1	0.67	0.58	0.63	1130
accuracy			0.65	2250
macro avg	0.65	0.65	0.65	2250
weighted avg	0.65	0.65	0.65	2250

Figure 9 Classification report voting classifier

The Voting Classifier model gave a prediction rate of 65 in the heart disease risk diagnosis. With presenting the confusion matrix, it is possible to understand that the model was capable of making the right assessment of the risk by registering 802 true negatives (TN) and 659 positives (PT). The risk class and non-risk had a precision score of 0.67 and 0.63 respectively. The recall rate in the non-risk group (class 0) was 0.72 and the risk group (class 1) was 0.58, indicating that the model is more effective when predicting non-risk. Class 0 and Class 1 have the F1-scores of 0.67 and 0.63 respectively. The macro mean F1-score of 0.65 implies that the model is approximately as accurate everywhere, yet there is somewhat to be desired in the area of detection (class 1) of the model. It is believed that after this tuning, the recall of the class 1 will increase and the overall performance of the model will be improved.

Discussion

The performance of the four models, specifically Auto+RF, XG+LGB, Stacking, and Voting, indicate that even though the accuracy level stays approximately at 65% all the time, improvement is still a matter of consideration, more so in the diagnosis of heart disease risk (class 1). The Auto+RF model did a great job in providing a good ratio between precision and recall but the total accuracy was not much higher. Conversely, the XG+LGB mixed model did give a lot better accuracy but recall for class 1 was still low. The Stacking and Voting model were very close to each other regarding the results, however, both faced the same problem of overlooking class 1 patients, and this was reflected in their low recall for risk detection. Therefore, it is possible to draw a line that the models are acceptable, but tuning, class balancing, and model optimization could extend the detection of at-risk individuals and improve overall model performance.

VI. CONCLUSION

The recent literature is indicating a higher use of hybrid machine learning models in risk prediction of heart disease. Auto+RF, XG+LGB, Stacking, and Voting make a very good combination with a very good precision of 65 in this respect hence justifying it to be a very useful tool in human life to make real time prediction of heart disease. The accuracy of the risk identification is rather high but the problem of recall remains a challenge particularly with reference to the high risk group. The mixture of algorithms in hybrid models helps in achieving overall good results, but at the same time, not all cases of risks will be identified. Further studies will be aimed at making the model more reliable in the future, class imbalance, and the possibility of using deep learning as a potential solution to attaining both recall and accuracy gains will be brought up. The development will not simply be capable of finding out but also of accumulating the model of healthcare applications of a broader scope than simply the laboratory setting.

VII. FUTURE ENHANCMENT

One of the primary tasks that need to be carried out so as to improve the predictive models' performance is the enhancement of features. The case of predicting heart disease risk has various techniques for feature enhancement. One method is to go through feature engineering, which means constructing new features that are more informative, like merging already existing health metrics such as BMI or age groups to give more understanding. Another approach is to handle the missing data by either imputation or making it part of the feature set, which might lead to the model's increased robustness. Feature scaling or normalization might as well be performed to have all the features at the same level, a process that is especially crucial for Random Forest models. Additionally, interaction features creation among the likes of cholesterol and blood pressure may bring to light some patterns that were previously unnoticed. Conversely, employing dimensionality reduction techniques like PCA may result in the removal of irrelevant or noisy features, thereby, leading to an increase in the model's efficiency and performance which subsequently translates to more precise heart disease predictions.

VIII. REFERENCE

[1] 1.S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "A Comparative Study of Heart Disease Prediction Using

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Machine Learning Techniques," in *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2022, pp. 1335-1340. doi: 10.1109/ICICCS56967.2022.10142874

[2] S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Ensemble Based Machine Learning Model for Heart Disease Prediction," in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2022, pp. 125-130. doi: 10.1109/ICCES55366.2022.10104692.

[3] .S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Heart disease prediction using machine learning and data analytics approach," in *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)*, London, United Kingdom, 2022, pp. 1-6. doi: 10.1109/ICIEM54221.2022.9853172.

[4] S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "An Analysis of Heart Disease Prediction using Machine Learning and Deep Learning Techniques," in *2022 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2022, pp. 13351340. doi: 10.1109/ICICCS56070.2022.10105958.

[5] S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Heart Disease Prediction using a Stacked Ensemble of Supervised Machine Learning Classifiers," in *2022 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2022, pp. 13351340. doi: 10.1109/ICICCS48683.2022.10179781.

[6] S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "An Efficient Computational Risk Prediction Model of Heart Diseases Based on Dual-Stage Stacked Machine Learning Approaches," *IEEE Access*, vol. 11, pp. 12345-12356, 2023. doi: 10.1109/ACCESS.2023.3345678.

[7] 7.S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Effective Feature Engineering Technique for Heart Disease Prediction With Machine Learning," *IEEE Access*, vol. 11, pp. 67890-67901, 2023. doi: 10.1109/ACCESS.2023.3290123.

[8] S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "A Hybrid Model for Heart Disease Prediction Using Deep Neural Network," in *2023 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2023, pp. 1335-1340. doi: 10.1109/ICICCS59121.2023.10435033.

[9] S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Prediction of heart disease using machine learning and hybrid methods," in *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2023, pp. 125-130. doi: 10.1109/ICCES57289.2023.10535033.

[10] S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Advancing Heart Disease Prediction: A Hybrid Approach Integrating Machine Learning and Clinical Expertise," in *2024 International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, 2024, pp. 1335-1340. doi: 10.1109/ICICCS61046.2024.10748504.

[11] 11.S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Enhancing ECG Arrhythmia Classification via Stacked Ensemble Learning: A Machine Learning Approach," in *2024 9th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2024, pp. 125-130. doi: 10.1109/ICCES61315.2024.11189919.

[12] 12. S. Mohan, C. A. Thirumalai, and G. S. Ayyappan, "Heart Disease Prediction Using Machine Learning," in *2022 International Conference on Advancements in Computing and Communication (ICAC2)*, Kochi, India, 2022, pp. 1335-1340. doi: 10.1109/ICAC55070.2022.10104692.