Performance Comparison between Meta-classifier Algorithms for Heart Disease Classification

Nureen Afiqah Mohd Zaini, Mohd Khalid Awang
Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, 22000 Tembila, Terengganu, Malaysia

***Abstract*—The rise in heart disease among the general population is alarming. This is because cardiovascular disease is the leading cause of death, and several studies have been conducted to assist cardiologists in identifying the primary cause of heart disease. The classification accuracy of single classifiers utilised in most recent studies to predict heart disease is quite low. The accuracy of classification can be enhanced by integrating the output of multiple classifiers in an ensemble technique. Even though they can deliver the best classification accuracy, the existing ensemble approaches that integrate all classifiers are quite resource-intensive. This study thus proposes a stacking ensemble that selects the optimal subset of classifiers to produce meta-classifiers. In addition, the research compares the effectiveness of several meta-classifiers to further enhance classification. There are ten types of algorithms, including logistic regression (LR), support vector classifier (SVC), random forest (RF), extra tree classifier (ETC), naïve bayes (NB), extreme gradient boosting (XGB), decision tree (DT), k-nearest neighbor (KNN), multilayer perceptron (MLP), and stochastic gradient descent (SGD) are used as a base classifier. The construction of the meta-classifier utilised three different algorithms consisting of LR, MLP, and SVC. The prediction results from the base classifier are then used as input for the stacking ensemble. The study demonstrates that using a stacking ensemble performs better than any other single algorithm in the base classifier. The meta-classifier of logistic regression yielded 90.16% results which is better than any base classifiers. In conclusion, we could assume that the ensemble stacking approach can be considered an additional means of achieving better accuracy and has improved the performance of the classification.**

***Keywords*—Heart disease prediction; ensemble stacking; multi classifier**

I. Introduction

According to research published by the World Health Organization (WHO), heart disease has been one of the leading causes of death worldwide. It is estimated to have reached 17.5 million in 2012, 17.9 million in 2016, and 22.2 million by 2030 [1]. Heart disease is very dangerous as it can be a silent killer to a patient. The individual must be alert to the main symptom of heart disease. The usual symptoms of heart disease risk factors include tobacco use, alcohol consumption, physical activity, fruit and vegetable consumption, hypertension, and obesity [2]. Providing high-quality care at affordable rates is a significant problem for the healthcare industry, and accurate diagnoses are the hallmarks of high-quality care [3]. Data mining is an interdisciplinary subject of computer science and statistic whose overall objective is to extract information from a dataset and transform the data into a usable structure [3], [4].

Using magnetic resonance imaging (MRI) and CT angiography to diagnose heart disease is highly complex. The equipment is bulky and prohibitively expensive for most individuals [5]. This research proposes a technique that can reduce the time and effort required for a specialist to diagnose cardiac problems in patients. The study of how to build compelling multi-classifier ensembles has been one of the most studied, and in most cases, ensemble techniques perform better than single classifiers [6]. The approach presented in this paper uses the stacking ensemble method to enhance the accuracy of heart disease prediction. This method is used to improve the overall accuracy of predictions, as the combination of models provides a collaboratively general solid model [7].

The single classifier approach is used most of the time in heart disease prediction research. The accuracy of classification may be improved by combining the results of many different classifiers using a method known as ensemble classification. Despite having the best classification accuracy, current ensemble approaches that include all classifiers are very resource-intensive. In order to create meta-classifiers, the stacking ensemble utilized in this study finds the best subset of classifiers. The study also looks at how effective various meta-classifiers are in improving classification.

Section II of this paper presents a review of related literature. Section III covers the research methodology that intricates the details of datasets and machine learning techniques. The overall results of the experiment are then presented in Section IV. Finally, a conclusion is presented in Section V.

II. Literature Review

A. Data Mining

Data mining is a cognitive method that identifies hidden approach patterns in massive datasets [8]. The effectiveness of data mining depends heavily on the method employed, and the feature is chosen since there are duplications and inconsistencies in healthcare industry medical datasets [9]. Data mining can be applied to improve disease risk assessment, intervention design, and monitoring of chronic conditions [10]. Hence, a reduction in patient admissions and insurance claims is possible.

Data mining is seen as critical yet challenging work that must be completed precisely and efficiently, especially in the healthcare environment. Comparison of different types of classification in machine learning has the potential to provide

high classification accuracy. Prediction accuracy may vary depending on the learning technique used [11].

The classification, which is one type of data mining activity, is crucial since it has the ability to classify such as identify the two categories whether the patient diagnoses with heart disease or non-heart disease. Several usual symptoms can be linked to detect heart disease, but not all are accurate. On the other hand, machine learning, a subset of artificial intelligence, is widely used in renowned research and can be utilised in the healthcare industry to overcome this issue. In return, it can help an expert to detect the disease and suggest an appropriate action to be taken.

B. Ensemble Method

A single classifier, such as Neural Network or Decision Tree, has been widely used in heart disease classification, but the performance of such algorithms is still lacking and need to be improved. Meanwhile, compared to a single classifier, the accuracy of ensemble methods such as stacking is significantly greater. Ensemble method approaches frequently increase predictive performance and this is due to various factors. Overfitting avoidance, computational advantage, and representation are the main factors why the ensemble method is able to improve predictive performance [12], [13]. To distinguish the main factors of the ensemble method, the following description provides examples of both:

1) Overfitting avoidance: Overfitting happens when the classifiers are tightly matched to the training data and thus lose the capacity to generalise to the test data. As a result, the performance of classification in the test set is inferior to the training set [14].

2) Computational advantage: Individual learners who undertake local searches may become mired in local optimality. By merging several learners, the ensemble method reduces the likelihood of achieving a local minimum [12].

3) Class imbalance: When there is a class imbalance in training data, learners will often overclassify the majority group because of its higher prior probability. Consequently, occurrences belonging to the minority group are more frequently misclassified than instances belonging to the majority group [15].

4) Curse of dimensionality: For high-dimension datasets, dimensionality reduction is performed prior to applying the algorithm in order to minimise the impacts of the curse of dimensionality. The more features of the dataset, the more challenges are produced for the algorithm used [16].

This study purposely proofs the concept of learning and making decisions is not new to humans. As a human, we employ this concept on a regular basis to make significant decisions, such as seeking advice from multiple experts or consulting with multiple physicians before undergoing major medical treatment [17].

C. Ensemble Stacking

Boosting, bagging, and stacking are the three most prominent ensemble learning approaches in machine learning. Ensemble techniques have been utilised in the current study to improve heart disease prediction classification accuracy. When weak learners are combined with meta-learners in tandem, a more accurate prediction can be made.

Stacking often accounts for a variety of weak learners [18]. Stacking architecture improves the accuracy of classification over a single classifier as it uses various ways to solve classification problems [19]. Stacking is a learning strategy that uses a meta-classifier to integrate the results of numerous basic classifiers learnt on the same dataset. Each algorithm has its own set of benefits and drawbacks. By merging the classifiers, the outcome may be enhanced.

D. Classification Algorithm in Data Mining

There are two main data mining approaches which are descriptive and predictive. These approaches uncover a hidden pattern in the vast data [20]. Classification can be categorised as supervised machine learning, in which we must train the data. Functional representations include decision trees, logistic regression, naïve bayes, extreme gradient boosting, k-nearest neighbor, extra tree classifier, multilayer perceptron, stochastic gradient classifier, random forest, and support vector machine.

A health management system is often used inside the healthcare setting to keep track of the large amounts of data pertaining to patients and the treatment they get. Such knowledge is useful, particularly if we put it to use in real-world situations. The majority of the medical database consists of discrete data. Thus, making decisions based on discrete facts becomes complex and difficult. Machine Learning which is a subset of data mining, handles large scale well-formatted datasets efficiently [21].

Numerous strategies for predicting heart disease are proposed, each of which employs a unique set of techniques and algorithms. Gaining great service at an affordable price remains the prime and challenging problem for healthcare organisations [22]. This study aims to help the specialist in the detection of heart disease in patients in the early stages.

E. Data Mining in Heart Disease Prediction

The use of data mining is an effort to meet the urgent need to extract relevant knowledge buried in clinical data, specifically to design a solution capable of predicting the presence or absence of heart disease using data mining [23]. In predicting heart disease, several data mining techniques such as regression, clustering, association rule, and classification algorithms consists of naïve bayes, decision tree, random forest, and k-nearest neighbor to categorise various heart disease [24].

According to [25], accuracy and receiver operating characteristics (ROC) are compared. The researcher obtained 303 records with 14 sets of variables and divided the data into training and testing. ROC plot is the most used for evaluating classifier performance. It is based on specificity and sensitivity, which are specificity measures of the negative part, and sensitivity is based on the positive part. Four machine learning algorithms, consisting of logistic regression, random forest, stochastic gradient boosting and support vector machine, are tested for this research. As a result, a comparison between the accuracy and ROC of the model prediction shows

that logistic regression is the highest prediction, with an accuracy of 87.00%.

The research by [26] used Gini index and prediction models. Neural network and Gini index are tested, and the results performed that Gini index results with the most remarkable precision. A decision tree is used to predict the accuracy and sensitivity of coronary illness. Voting methods are known to produce a more precise decision tree. The decision tree makes up of these critical elements, which are the decision hub and edge or branch. The decision hub determines the test features, while the edge or chapter compares to one of the possible quality esteems and denotes one of the property findings. The class to which the question belongs is contained in a leaf, also known as an answer hub.

The author of the study [27] compared the classifier using principal component analysis (PCA) and without PCA. The authors experimented with five types of algorithms for the research. Decision tree, logistic regression, support vector machine, multilayer perceptron, and naïve bayes are tested. Logistic regression achieved the highest with 86.00% without PCA, and random forest achieved the highest with a PCA of 84.00% accuracy. The following areas of interest have seen the most application of PCA, such as classification, data clustering, and dimension reduction.

The study by [28] used a dataset from UCI and proposed four classifiers for the research. Naïve bayes, decision trees, support vector machines, and k-nearest neighbour are tested. According to the researcher, naïve bayes achieves the highest accuracy with 88.67%. Naïve bayes is a straightforward algorithm to implement. Asides, this algorithm also can be categorised as space efficient and fast to train. The author claims that knowledge discovery is essential to handling a large dataset for the identification of heart disease. Managing a massive dataset for heart disease detection and finding the relevant information to forecast a heart attack in its early stages based on patient indicators is challenging.

In 2020, [29] analysed the detection of heart disease using these six algorithms: artificial neural network (ANN), logistic regression, k-nearest neighbour, support vector machine, decision tree, and naïve bayes. The researchers also stated the standard state-of-the-art features selection algorithms consisting of Relief, MRMR, LASSO, and Local-learning-based-features selection (LLBFS) are used to select the features. The researcher also proposes fast conditional mutual information (FCMIM) features selection algorithms for feature selection. LOSO technique (Leave-one-subject-out) cross-validation is applied to select the best hyper-parameter for the best model selection. The best accuracy achieved is SVM with the proposed feature selection algorithm (FCMIM) of 92.37%, which is very good compared to the previously proposed method.

Based on the above discussion, there are two approaches to heart disease prediction: single classifier and meta-classifier. Most of the heart disease prediction research uses the single classifier approach. Existing ensemble methods that incorporate all classifiers are highly resource-intensive, notwithstanding their ability to provide the highest classification accuracy. This paper thus presents a stacking ensemble that picks the optimum subset of classifiers in order to generate meta-classifiers. In addition, the study examines the efficacy of many meta-classifiers to improve classification. The following section will describe our proposed stacking ensemble in detail.

III. Methodology

This study is based on the UCI dataset of people with cardiac disease. There are 303 records and 13 attributes in total, which are divided into training and test sets [33]. Table I explains the dataset features for the heart-disease patients.

TABLE I. Description of Attributes Dataset Features

No	Feature Name	Type of data	Data Description
1	X_1= age	Numeric	The age of the patients
2	X_2=sex	Nominal	Female, male
3	X_3= cp	Nominal	Chest pain type Value 0 = Typical angina Value 1 = Atypical angina Value 2 = non-anginal pain Value 3 = asymptomatic
4	X_4= trestbps	Numeric	Resting blood pressure
5	X_5 =Chol	Numeric	Serum cholesterol
6	X_6 = Fbs	Nominal	Fasting blood sugar
7	X_7=restecg	Nominal	Resting electrocardiographic
8	X_8=thalach	Numeric	Max heart rate
9	X_9 = exang	Nominal	Exercise-induced angina
10	X_{10}=oldpeak	Nominal	ST depression induced by exercise relative to the rest
11	X_{11} =Slope	Nominal	The slope of peak exercise
12	X_{12} =Ca	Nominal	Number of major vessels
13	X_{13} = Thal	Nominal	The heart status
14	Y = Target	Nominal	Diagnosis of heart disease

A. Proposed Stacking Ensemble

Fig. 1 is a framework on how our proposed meta-classifier is designed, and Table I explains the details of the dataset. The goal of the research is to choose the optimal subset of base classifiers for creating the meta-classifier. Further, the best combination of base classifiers and meta-classifiers is determined by comparing several meta-classifier learning algorithms. The proposed stacking model is described in more detail below:

1) Split the data into training and test set.

2) Develop base classifiers based on ten different learning algorithms.

a) Train and test the base classifiers.

b) Rank the base classifiers based on the accuracy of prediction performance.

3) Develop meta-classifier based on three different learning algorithms.

a) Select the optimum number of base classifiers (2, 4, 5, 7 and 10).

b) Test the performance of the meta-classifiers.

4) The final model of ensemble stacking is obtained from the combination of the optimum subset of base classifiers and the meta-classifier algorithm.

From the UCI dataset, 13 attributes and ten base classifiers are utilised, as shown in Fig. 1. The output from the level-1 base classifiers is then applied to the level-2 meta-classifier and followed by the selection of optimal subset of base classifiers. Finally the model is evaluated using three distinct meta-classifiers learning algorithms.

B. Level-1 base Classifiers Construction

The base model classifier is stimulated by certain training data and generates different predictions. The main stage of developing the ensemble technique is to populate the database with a group of base classifiers. As a result, each of the base models will produce its own set of predictions. Base classifiers for this research are created using ten different learning algorithms. Listed below are the selected algorithms:

1) C_1 = LogisticRegression()
2) C_2 = KNeighborsClassifier()
3) C_3 = DecisionTreeClassifier()
4) C_4 = RandomForestClassifier()
5) C_5 = GaussianNB()
6) C_6 = XGBoostClassifier()
7) C_7 = SVC ()
8) C_8 = MLPClassifier ()
9) C_9 = SGDClassifier()
10) C_{10} = ExtraTreesClassifier()

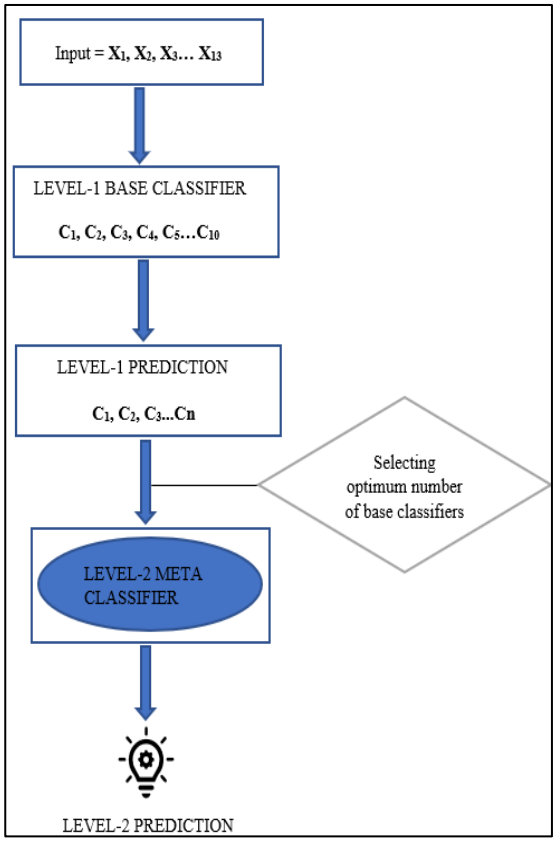

Fig. 1. Proposed Ensemble Stacking for Heart Disease Classification.

C. Level-2 Meta-classifiers Construction

Results obtained from the base classifier are then tested for the ensemble stacking method. Three different learning algorithms are tested with different subset of base classifiers. Meta-learning encompasses a wide range of activities, such as observing the performance of different machine learning models about learning tasks, learning from metadata, and performing a faster learning process for new tasks. A few learning algorithms' classification performances will be evaluated to select optimal meta-model classification. The selected learning algorithms are described below:

1) Meta-Classifier using Logistic Regression ()
2) Meta-Classifier using Support Vector Machine ()
3) Meta-Classifier using Multi-Layer Perceptron ()

Based on their performance, the study has ranked the base classifiers. Five alternative subsets, including 2, 4, 5, 7, and 10 classifiers, were examined using various combinations of the base classifiers. The subset of 5 classifiers is selected as input to the meta-classifier based on the experiments since it outperforms other combinations. There are three types of learning algorithms considered in this study. First, the meta-model is based on Logistic Regression. The second is based on a Support Vector Machine classifier, and the third is based on a Multilayer Perceptron. The performance of each learning algorithm is compared using performance measurements, as discussed in the next section.

D. Performance Measurement

Performance measurement is crucial to data mining evaluation. Several indicators, including accuracy, precision, recall, and F1-score, are used to evaluate the prediction rate and determine the validity of the models [30]. In this research, accuracy is chosen compared to the area under the curve (AUC) because of the percentage obtained. Precision is an upbeat class in which the model is predicted correctly or correctly outputs supplied by the model. The following formula can be used to compute as below [31]:

$$\text{Precision} = TP / (TP+FP) \quad (1)$$

Recall must be high as possible and can be concluded as how the model is predicted correctly. Below is how it can be computed [31]:

$$\text{Recall} = TP / (TP+FN) \quad (2)$$

A model with low precision, high recall, or high precision, low recall is hard to compare. To solve the problem, F-score can be used. This score allows for assessing both recall and precision simultaneously. The following formula can be used to compute [32]:

$$\text{F-Measure} = 2*\text{Recall}*\text{Precision} /(\text{Recall}+\text{Precision}) \quad (3)$$

IV. Results and Discussion

The accuracy of heart disease classification will be discussed in this chapter and the findings of a series of experiments undertaken to validate the classification accuracy performance. The outcomes of this study are compared and analysed depending on the algorithm used to test the performance of the proposed strategy.

The performance measures include receiver operating characteristic area under the curve (ROC AUC), precision, recall, and F1 measure. Table II shows the classification's average performance based on precision, accuracy, recall, AUC, and F1 measure.

Table II shows the experimented results for base classifiers based on the accuracy, precision, recall, and F1-score. We have tested and compared ten base models' classifiers, logistic regression, random forest, k-nearest neighbour, decision tree, naïve bayes, support vector machine, extreme gradient boosting, multilayer perceptron, stochastic gradient descent, and extra tree classifier. The best model is MLP, with an accuracy of 88.52%, precision of 86.48%, recall of 94.11%, F1-score of 90.14%, and AUC of 87.79%. While when compared with the decision tree, the accuracy achieved is 72.13%, which is the lowest amongst others, with the precision of 79.31%, recall of 67.64%, F1-score of 73.01%, and AUC of 72.71%.

Table III shows that logistic regression, support vector classifier and multilayer perceptron are chosen as the meta-classifiers. These three meta-classifiers are then experimented with using the subset of base-classifiers. The combination of the base classifier is in the group of two, four, five, seven, and ten classifiers. The formation of the subset is based on the accuracy performance of each of the base classifiers. The base classifiers are ranked according to their prediction performance. Logistic regression is recorded as the best algorithm, tested on a subset of five base classifiers with an accuracy of 90.16%. SVC and MLP achieved 83.60% and 88.52% using the five classifiers.

TABLE II. DETAILS OF PERFORMANCE MEASURE OF THE CLASSIFICATION FOR THE BASE MODEL

Base Model	Accuracy	Precision	Recall	F1-score	AUC
Logistic Regression	85.24%	83.78%	91.17%	87.32%	84.47%
Random Forest	83.60%	85.29%	85.29%	85.29%	83.38%
KNN	81.96%	82.85%	85.29%	84.05%	81.53%
Decision Tree	72.13%	79.31%	67.64%	73.01%	72.71%
Naïve Bayes	85.24%	83.78%	91.17%	87.32%	84.47%
SVC	86.88%	84.21%	94.11%	88.88%	85.94%
XGB	85.24%	85.71%	88.23%	86.95%	84.85%
MLP	88.52%	86.48%	94.11%	90.14%	87.79%
SGD Classifier	83.60%	80.00%	94.11%	86.48%	82.24%
Extra Tree Classifier	86.88%	88.23%	88.23%	88.23%	86.71%

TABLE III. SUMMARY OF META-MODEL ACCURACY USING A DIFFERENT SUBSET OF CLASSIFIERS

Meta-Model / Classifiers	Subset of 2	Subset of 4	Subset of 5	Subset of 7	Subset of 10
Logistic Regression	86.88 %	88.52%	90.16%	86.88%	83.60%
SVC	86.88 %	85.24%	83.60%	83.60%	83.60%
MLP	86.88 %	88.52%	88.52%	86.88%	86.88%

Based on the experiment, the subset of two base classifiers, LR, SVC, and MLP achieved the same accuracy of 86.88%. LR and MLP resulted in the same accuracy of 88.52% using a subset of four classifiers, and SVC achieved 85.24%. For a subset of seven classifiers, again, LR and MLP achieved the same accuracy of 86.88%, and SVC achieved 83.60% of accuracy. This experiment also experimented using a subset of ten classifiers. The accuracy achieved for LR and SVC is 83.60%, and MLP attained an accuracy of 86.88%.

As a result, the ensemble technique showed that the best accuracy was achieved utilising a subset of five classifiers using logistic regression as meta-model classifiers. The presumption is accurate when comparing LR to SVC and MLP. This study shows that both the utilisation of large subsets and a limited number of subsets is undesirable. According to the experiments, it's crucial to identify the best subset of base classifiers before building the meta-classifiers in ensemble stacking since doing so could enhance classification performance while also making efficient use of the available resources.

V. CONCLUSION

In conclusion, many people might prevent chronic heart disease attacks with early detection of heart disease. Most research uses a single classifier for prediction, yet the outcome is still unsatisfactory. While compared to utilising a single base classifier, the accuracy performance of heart disease in this research has improved when employing the meta-classifier, such as stacking for the ensemble technique.

Meta-classifier of the Logistic Regression algorithm and a combination of five subsets of base classifiers has achieved the highest accuracy compared to others, with 90.16% accuracy. The subset of base classifiers consists of support vector machine (SVM), multilayer perceptron (MLP), extra tree classifier (ETC), naïve bayes (NB) and extreme gradient boosting (XGB). This study can be proposed to the specialist for the future improvement of heart disease prediction and can be used for medical purposes. The use of ensemble methods such as stacking can enhance the prediction of accuracy.

REFERENCES

[1] G. İ. Bayram Şahin, “Risk factors of deaths related to cardiovascular diseases in World Health Organization (WHO) member countries,” Heal. Soc. Care, vol. 30, no. 1, pp. 73–80, 2020.

[2] Y. Ruan et al., "Cardiovascular disease (CVD) and associated risk factors among older adults in six low-and middle-income countries: results from SAGE Wave 1," BMC Public Health, vol. 18, no. 778, pp. 1–13, 2018.

[3] S. Anitha and N. Sridevi, "HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES," J. Anal. Comput., vol. 13, no. 2, 2019.

[4] M. Mohamad, M. Y. M. Saman, and N. A. Hamid, "Complexity Approximation of Classification Task for Large Dataset Ensemble Artificial Neural Networks," Lect. Notes Electr. Eng., vol. 520, no. April, pp. 195–202, 2019.

[5] R. Hasan, S. Ahmed, and A. Afridi, "Heart Sound Analysis for Detection of Cardiovascular Diseases (CVDs) Through Segmentation," Int. J. Sci. Eng. Res., vol. 11, no. Jun 2020, pp. 1–6, 2021.

[6] T. G. Dietterich, "Ensemble methods in machine learning," Lect. Notes Comput. Sci., vol. 1857, pp. 1–15, 2000.

[7] R. Atallah and A. Al-Mousa, "Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method," 2019 2nd Int. Conf. New Trends Comput. Sci. ICTCS 2019 - Proc., pp. 1–6, 2019.

[8] M. Tarawneh and O. Embarak, Hybrid Approach for Heart Disease Prediction Using Data Mining Techniques, vol. 29. Springer International Publishing, 2019.

[9] M. S. Amin, Y. K. Chiam, and K. D. Varathan, "Identification of significant features and data mining techniques in predicting heart disease," Telemat. Informatics, vol. 36, pp. 82–93, 2019.

[10] P. Perner, "Advances in Data Mining Applications," Lect. Notes Eng. Comput. Sci., vol. 5, no. 3, pp. 1–6, 2018.

[11] A. K. Dwivedi, "Performance evaluation of different machine learning techniques for prediction of heart disease," Neural Comput. Appl., vol. 29, no. 10, pp. 685–693, 2018.

[12] O. Sagi and L. Rokach, "Ensemble learning: A survey," Wiley Interdiscip. Rev. Data Min. Knowl. Discov., vol. 8, no. 4, pp. 1–18, 2018.

[13] S. Hisham, M. Makhtar, and A. A. Aziz, "Combining Multiple Classifiers using Ensemble Method for Anomaly Detection in Blockchain Networks: A Comprehensive Review," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 8, pp. 404–422, 2022.

[14] M. S. Santos, J. P. Soares, P. H. Abreu, H. Araujo, and J. Santos, "Cross-validation for imbalanced datasets: Avoiding overoptimistic and overfitting approaches [Research Frontier]," IEEE Comput. Intell. Mag., vol. 13, no. 4, pp. 59–76, 2018.

[15] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," J. Big Data, vol. 6, no. 1, 2019.

[16] V. Kirankumar, S. Ramasubbareddy, G. Kannayaram, and K. N. Kumar, "Classification of heart disease using support vector machine," J. Comput. Theor. Nanosci., vol. 16, no. 5–6, pp. 2623–2627, 2019.

[17] K. Raza, Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule. Elsevier Inc., 2019.

[18] M. K. Awang, M. Makhtar, N. Udin, and N. F. Mansor, "Improving Customer Churn Classification with Ensemble Stacking Method," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 11, pp. 277–285, 2021.

[19] A. El Sheikh, N. Mahmoud, and A. Keshk, "Heart Disease Classification Based on Hybrid Ensemble Stacking Technique," IJCI. Int. J. Comput. Inf., vol. 8, no. 2, pp. 1–8, 2021.

[20] S. Umadevi and K. S. J. Marseline, "A survey on data mining classification algorithms," Proc. IEEE Int. Conf. Signal Process. Commun. ICSPC 2017, vol. 2018-Janua, no. July, pp. 264–268, 2018.

[21] I. G. J. Sree, D. Prasanna, B. Y. Rachana, and N. J. Basha, "Performance Analysis of Heart Disease Diagnosis Using Machine," IJCST, vol. 9, no. 5, pp. 1–6, 2021.

[22] A. N. Repaka, S. D. Ravikanti, and R. G. Franklin, "Design and implementing heart disease prediction using naives Bayesian," Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019, vol. 2019-April, no. Icoei, pp. 292–297, 2019.

[23] M. Bárbara, M., Ferreira, D., Cristiana, N., António, A., & José, "Data Mining for Cardiovascular Disease Prediction," J. Med. Syst., vol. 1, no. 45, pp. 1–8, 2021.

[24] S. Guruprasad, V. L. Mathias, and W. Dcunha, "Heart Disease Prediction Using Machine Learning Techniques," 2021 5th Int. Conf. Electr. Electron. Commun. Comput. Technol. Optim. Tech. ICEECCOT 2021 - Proc., vol. 1, no. 6, pp. 762–766, 2021.

[25] N. B. Muppalaneni, M. Ma, and S. Gurumoorthy, Soft Computing and Medical Bioinformatics. Springer Singapore, 2019.

[26] K. Mathan, P. M. Kumar, P. Panchatcharam, G. Manogaran, and R. Varadharajan, "A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease," Des. Autom. Embed. Syst., vol. 22, no. 3, pp. 225–242, Sep. 2018.

[27] C. S. Wu, M. Badshah, and V. Bhagwat, "Heart disease prediction using data mining techniques," in ACM International Conference Proceeding Series, 2019, pp. 7–11.

[28] S. Nayak, M. K. Gourisaria, M. Pandey, and S. S. Rautaray, "Prediction of heart disease by mining frequent items and classification techniques," 2019 Int. Conf. Intell. Comput. Control Syst. ICCS 2019, no. Iciccs, pp. 607–611, 2019.

[29] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan, and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," IEEE Access, vol. 8, no. Ml, pp. 107562–107582, 2020.

[30] S. Guruprasad, V. L. Mathias, and W. Dcunha, "Heart Disease Prediction Using Machine Learning Techniques," in 2021 5th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques, ICEECCOT 2021 - Proceedings, 2021, vol. 1, pp. 762–766.

[31] E. Nikookar, "Hybrid Ensemble Framework for Heart Disease Detection and Prediction," IJACSA, vol. 9, no. 5, pp. 243–248, 2018.

[32] H. Khdair and N. M. Dasari, "Exploring Machine Learning Techniques for Coronary Heart Disease Prediction," Int. J. Adv. Comput. Sci. Appl., vol. 12, no. 5, pp. 28–36, 2021.

[33] D. Newman D, S. Hettich, C. Blake and C. Merz, "UCI Repository of machine learning databases." http://www.ics.uci.edu/~mlearn/ MLRepository.html, 2021.