Combining A Hybrid Genetic Algorithm with A Fuzzy Logic Classifier Enhances Heart Disease Diagnosis

S KOKILA [1], S.PRAVEENA [2]

[1] Department of Computer Science, Salem kongunaadu Arts and science college, Mamangam -636302.
kokila333@gmail.com
[2] Department of computer science
Salem kongunaadu arts and science college,mamangam,Salem

ABSTRACT

Because cardio illness has a significant impact on death rates worldwide, diagnosing heart disease is an essential field of healthcare. The study proposes a unique method, the HDD-GA-FL model, to enhance Heart Disease Diagnosis (HDD) that combines a hybrid genetic algorithm (GA) along with a fuzzy logic classifier (FL). The suggested hybrid system aims to overcome the difficulties brought on by the intricacy and ambiguity involved in diagnosing cardiac disease. Fuzzy logic classifiers are used to analyze ambiguous medical data, while genetic algorithms are used for choosing features and optimization. Combining these two methods provides a strong foundation for precise and effective diagnosis. Experiments on an extensive dataset with different clinical factors and cases of heart disease are conducted to assess the efficacy of the hybrid strategy. Compared to conventional diagnostic techniques, there have been significant improvements in diagnostic reliability and accuracy. When navigating the complex feature space involved in diagnosing heart illness, the combination of GA and FLC performs better than alone. It can capture minute associations and trends in the data. The suggested hybrid system has great potential for real-world application in clinical settings, providing doctors with an invaluable instrument for accurately diagnosing and detecting cardiac problems early on. This method advances the latest developments in cardiac healthcare by utilizing the complementary strengths of fuzzy logic classifiers and genetic algorithms, eventually improving patient outcomes and lowering healthcare costs.

Keywords: Heart disease diagnosis; Hybrid genetic algorithm; Fuzzy-logic classifier; Hybrid system; Feature selection; Optimization.

1. Introduction

Recent developments in machine learning and computational intelligence have opened up exciting new possibilities for improving the diagnosis of cardiac disease. Using data-driven analysis, these methods can conclude relevant relationships and trends from large datasets, potentially leading to more objective, effective, and precise diagnostic results. In general, angiography is seen to be the most effective method for detecting coronary vein infections, but it has a lot of adverse effects and is very expensive. As a result, many researchers have been told to use information mining and machine learning to find elective modes. As a result, the study has provided an accurate hybrid approach method for diagnosing cardiovascular disease (CVD) in this process [1]. The primary concern in data mining and artificial intelligence is finding an appropriate data representation from all features. Not every original feature is helpful for regression or classification tasks [2]. Heart disease carries several hazards, including those

related to gender, age, obesity, hypertension, high cholesterol levels, diabetes, smoking, alcohol consumption, and family history [3]. The classification technique is one popular application of Machine Learning (ML) for healthcare diagnosis and forecasting. Classification accuracy usually gauges model performance [4]. The evolution of machine learning involves one crucial step: feature selection. Several techniques are used in feature selection, such as filtering, wrappers, and embedding. The hybrid method is one way to develop feature selection by combining many techniques. The goal of the hybrid approach is to have better characteristics than just one method applied [5]. Better accuracy requires optimisation, and while numerous methods exist, feature selection with genetic algorithms (GA) is one of the most sophisticated [6].

The suggested method combines machine learning, genetic algorithms, and statistical analysis techniques to choose an ideal subset from the entire feature space. This allows it to benefit from packing and filtering techniques [7]. Several researchers have reported better evaluation results by combining the fuzzy classifier with a successful fuzzy selection technique, creating a new hybrid model. This hybrid approach has demonstrated medical potential, especially in diagnosing illnesses [8]. To fine-tune the fuzzy rules, apply the fuzzification procedure to input variable values. The activated fuzzy rules help ascertain the collection of fuzzy outcomes. A certain level is considered into account for every fuzzy rule that is accessible. In this case, it produces the relevant fuzzy rules for illness control [9]. A neural fuzzy hybrid system based on genetic algorithms combines the computational analysis and design of one model created to predict cardiovascular illness. It examines an integrated system using neural networks, genetic algorithms, and fuzzy inference [10]. Because several diseases share symptoms, diagnosing them can occasionally be difficult, necessitating optimising approaches [11]. Techniques such as metaheuristics and algorithms for optimisation are frequently employed to provide intelligent diagnostic systems that improve classification accuracy.

2. Literature review

Li et al. [12] proposed that to reduce healthcare pressure during COVID-19, Artificial Intelligence (AI) is required to detect diabetes diagnosis. Particle swarm optimisation using K-means, Harmony Searching, Genetic Algorithm, and K-nearest neighbour classification are the three steps in this approach. The methodology outperforms previous methods with an accuracy of 91.65%. There are still generalizability, dataset variances, and ethical issues for broader implementation in healthcare settings. Taylan et al. [13] proposed a machine learning (ML) paradigm that combines neural networks, adaptive neuro-fuzzy inference, and support vector regression. Regarding CVD risk variables, hybrid models perform better on actual data from hospitals than classic ML algorithms. Sensitivity analysis highlights factors such as age, cholesterol, and glucose levels. Notably, the adaptive neuro-fuzzy inference system (ANFIS) shows better promise for CVD classification with 96.56% prediction accuracy, while SVR comes in second at 91.95%. Due to different data architectures in other nations and institutions, it is imperative to anticipate heart problems early and accurately. Menshawi et al. [14] proposed a novel voting mechanism used across two layers in a hybrid framework that integrates numerous machine learning and deep learning techniques to reduce bias. The selection of features comes before classification, producing a 95.6% accuracy rate higher than single-model methods. Because of its versatility, the system may be retrained for various datasets, demonstrating its potential for significant healthcare applications.

To diagnose CVD, the primary cause of death worldwide, Ali et al. [15] offered an expert system based on fuzzy logic. It uses seven input attributes for accurate diagnosis and integrates modules for fuzzification, information base, inference, and

defuzzification. Extensive collaboration with medical specialists guides the development of a comprehensive knowledge base among IF-THEN rules. By outperforming current CVD detection systems with 98.08% accuracy using the Cleveland dataset, the web-based integration of the system improves accessibility and affordability. Heart and blood vessel diseases, together known as CVD, account for 18 million deaths worldwide each year. Prevention requires early detection and the right kind of care. Predicting cardiac illness may be possible with machine learning algorithms, essential for medical diagnostics. Manikandan et al. [16] used the Cleveland Clinic's Heart Disease Dataset to evaluate Boruta feature selection in logistic regression, decision trees, and support vector machine (SVM) techniques. Boruta enhances algorithm performance; at 88.52%, logistic regression yields the maximum accuracy. Parveen et al. [17] proposed a hybrid model that blends ANN (artificial neural networks) and advanced fuzzy approaches to predict cardiovascular disease. If fuzzy TOPSIS helps with disease classification, ANN predicts disease risk. The accuracy of predictions is improved through attribute weighting using the Analytic Hierarchy Process (AHP). High measures of accuracy (0.99), precision (0.98), specificity (0.978), The F-measure (0.981), and sensitivity (0.996) demonstrate the model's superiority over conventional approaches in comparison analysis. By combining fuzzy logic with artificial neural networks, this work seeks to improve knowledge-based systems' predictive and classification capabilities for cardiovascular disease.

Using the IEEE Dataport Heart Disease Dataset, Alanazi et al. [18] proposed Auto-WEKA, Decision Table/Naive Bayesian (DTNB), and Multiobjective Evolution (MOE) fuzzy classifier algorithms. Using hyperparameter and classifier selection optimisation, Auto-WEKA attains 100% accuracy. DTNB achieves 85.63% accuracy, and the MOE fuzzy classifier achieves 81.6%, representing room for improvement in precision and recall through more modification. Our findings support the continued improvement of classifier parameters and demonstrate the potential of machine learning in diagnosing CVD. Alshraideh et al. [19] proposed improving heart disease diagnosis and therapy. The study creates machine-learning models using the Jordan University Hospital's Heart Dataset (JUH). Particle swarm optimisation, or PSO, is investigated for feature selection with Random Forest, Support Vector Machines (SVM), Decision Trees, Naive Bayes, and k-nearest-neighbors (KNN). The suggested approach outperforms other algorithms with an accuracy of 94.3%; in particular, SVM with PSO is highlighted for better heart disease detection in Jordan. Reducing death rates from cardiovascular diseases (CVD) requires early detection. Kalaivani et al. [20] proposed feature selection techniques and machine learning models to find risk factors. Feature selection is improved using a hybrid method that combines the LASSO algorithm with Differential Entropy-based information gain. It enhances classification performance with Random Forest, improving accuracy, precision, and recall.

3. The proposed HDD-GA-FL framework

a. Dataset:

One of the most common chronic illnesses in the US, heart disease affects millions of people annually and has a significant financial impact on the nation's economy. About 647,000 people die from heart disease in the United States just each year, making it the most common cause of mortality [21]. Heart disease is occurring and has risk factors like diabetes, high blood pressure, chronic inflammation, ageing-related molecular changes, plaque accumulation inside more prominent coronary arteries, and high blood pressure. Although there are various forms of coronary artery disease, most people don't realise they have it until they have symptoms like chest discomfort, angina, or unexpected cardiac arrest. This information emphasises the value of screening programs and diagnostic procedures that can reliably identify heart disease in the

general population before adverse events, such as myocardial infarctions or heart attacks, occur. The Centres have recognised three significant risk factors of heart disease for the Control of Disease and Prevention: smoking, increased blood pressure, and increased blood cholesterol. All three risk factors are present in about half of the American population. The National Heart, Lung, as well as Blood Institute issues physicians with a more comprehensive list of variables to consider when diagnosing coronary heart disease, including age, sex, race or ethnicity, environment and occupation, lineage and genetics, lifestyle habits, and other medical conditions. An initial assessment of these typical risk factors, followed by blood and other tests, typically serves as the basis for the diagnosis.

Every year, the CDC collects information via a telephone health survey called the System for the Study of Behavioral Risk Factors (BRFSS). Every year, the study gathers data from over 400,000 Americans about chronic health conditions, disease-related risk factors, and the use of preventative treatments. From 1984 forward, it has been held annually. 253,680 questionnaire responses from the cleaned BRFSS 2015 are included in this dataset, which is mainly intended for use in binary heart disease categorisation. Not that this dataset exhibits a significant class disparity. While 23,893 respondents have experienced heart disease, 229,787 answers do not have it or have never had it.

Feature reduction improves classification performance while lowering computational costs. This research uses fuzzy logic classifiers to construct rule sets, and rough sets are employed for feature reduction to improve disease prediction outcomes. Hybrid Genetic Algorithm creates the solution set to obtain optimal rules for illness prediction. The steps in the illness prediction model are as follows: Attribute reduction, normalisation, and HGA-FL classification were applied to rough sets. First, the input dataset is normalised within the interval [0, 1]. The best attributes are chosen using the rough set-based technique. Two subsets of the decreased attributes will be created: training and testing datasets. HGA-FL receives the training dataset, while the suggested model is tested using the testing dataset. Figure 1 shows the proposed HDD-GA-FL framework. The following sub-sections are the description of each stage in the suggested model:
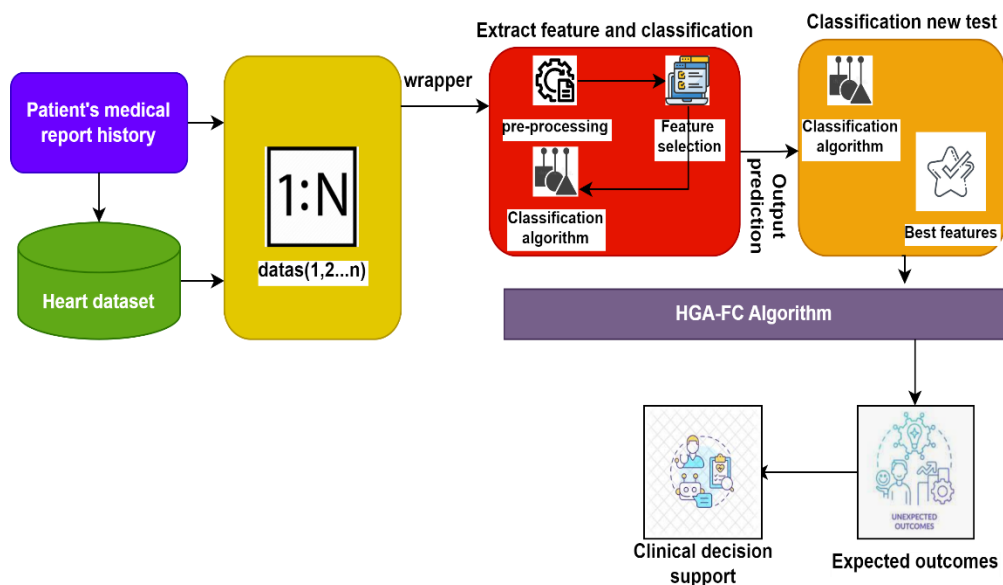

Figure 1: The proposed HDD-GA-FL framework

b. Hybrid Genetic algorithm

John Holland was the first to publish the genetic algorithm. This method is a specific type of evolutionary algorithm that finds approximate answers based on optimisation and search problems. It uses the heredity and mutation methods. By applying the process of species development, GA assists in resolving a wide range of ongoing issues. The "chromosome" significantly contributes to this method, comprising parameters with a single characteristic type. Each chromosome is made up of a collection of attributes. As shown in Figure 2, the flowchart conveys the workflow of the algorithm information's standard.
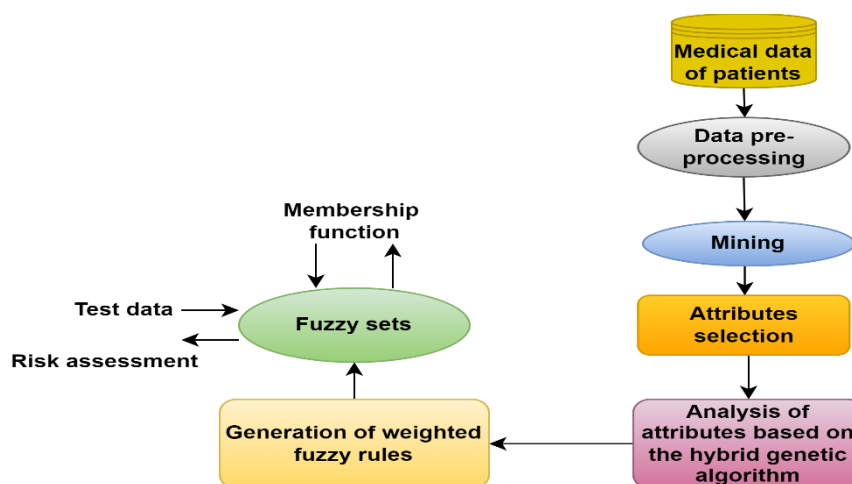

Figure 2. The GA-FL algorithm flowchart

A collection of these chromosomes frames a population. Following a sequence of focus, the chromosome obtained in the population, known as generations, answers the problem. Every genetic information is stored within the chromosomes. Every generation is better than the one before it because there is a growing chance of making the arrangement. Genetic processes like selection, mutation, cross-over, and fitness functions provide the solution in a whole generation.

Illustration of a chromosome

The crucial GA segment yields the best solution: the fitness function structuring. The fitness functions that are available in writing are typically subordinate. Specific optimal results for various information indices are not generated from several databases. The value of the fitness function estimates the proximity reached in the ideal solution.

The three distinct genetic operators are defined in the following section: The method used by the ultimate genetic operator is known as Selection. The mutation and crossover process selected parents from the designated population. The procedure also specifies the kind and number of appropriate parent chromosomes. In this instance, the likelihood corresponding to the chromosome's fitness rating will be used to choose it. In that technique, there's a greater possibility of the fitter chromosomes being chosen. The crossover method is used to select the parents from the population pool after the crossover point, and it also reverses the values' placements.

The values in the second portion of the enhanced chromosome will be substituted for those in the extra one, and the earlier settled values will be transferred from a single chromosome into the original component of the current chromosome (as shown in equation (1)).

$$Offspring = P1 + \alpha(P2 - P1) \tag{1}$$

Alpha (α) is a scaling factor that is continuously and arbitrarily selected, and P1 and P2 represent parent 1 and parent 2, respectively. A mutation arises from combining or changing the values of the gene in a parallel or genuine representation for each gene individually, based on the likelihood of the mutation value. The mutation administrator continuously searches and accelerates. It breaks free from around minima, and since it preserves a reasonable degree of population diversity, the optimal arrangement may result from providing an appropriate incentive for this behaviour.

The Gaussian shift is used in this analysis, and the values for scaling and contracting are respectively 0.05 and 1. This technique results in a higher value for the chromosome selected based on the zero-fixed Gaussian dispersion. The standard variance between generations is influenced by the therapist's parameter, which is a control parameter. The value of the scale establishes the standard deviation based on the underlying age. The different halting conditions involve the following: the procedure reaching its maximum number of generations; the client-set maximum time limit; a scenario in which individual health characteristics remain relatively unchanged for a pre-specified number of generations; and an instance in which the intended job achievement is not appreciably improved. The terms "stall future generations" and "stop-limit" describe these circumstances.

c. Fuzzy logic classifier

The work's main objective is to classify using fuzzy logic. This classification technique is crucial for artificial intelligence, especially for intermediate input value issues. This logic approach handles partial truth notation and provides approximate solutions; however, it does not yield an accurate result. The true values are categorised based on two factors: totally false and entirely true.

Pre-processing of data

Pre-processing information aims to eliminate user data from cardiac disease databases. Thus, data must be converted into configuration to forecast risk level. Because there is insignificant data on heart illness, the first rough information utilised in the expectation approach cannot be used directly. Instead, it must be cleaned, analysed, and changed during the information pre-handling step.

Mining of selected attributes

It is crucial to utilize the ongoing categorization of features relevant to the complete attribute shown in the dh_i datasets throughout this phase. This will enable the recurrence of each letter kind in the c_j class to be determined through database analysis. The attribute recurrence class for the whole dataset is ascertained by processing the quality classification recurrence. For constant features, get the recurrence by dividing those with similar widths. In this example, the recurrence order of the quality-based attribute classification within the class c_j is utilized, and the database requirements are determined by the use of heredity computation for mining the one-duration asset classification. Following sorting, the olist orderis utilized to identify several attribute groupings with the lowest level of support for each attribute. Two vectors, and $V_j(\text{max})$ and $V_j(\text{min})$, per every category are then subjected to the selected attribute categorization. Specifically, every attribute quality classification has a base value located in a single vector, and each quality's characteristic classification is contrasted in a second vector that contains the most extreme value.

$$V_j(\max) = \{\beta_{max}^1, \beta_{max}^2, \dots\dots, \beta_{max}^m\} \qquad (2)$$

Selection of suitable attributes

Using $V_j(\min)$, and $V_j(\max)$, two vectors, the relevant characteristics that were improved upon in the prior advancement are identified. The rationale for this advancement is that, although the information includes more features, none are particularly effective at predicting the persistent cardiac hazard. Detecting suitable properties should verify greater accuracy in the expectation of danger level. The deviation technique, which uses the quality subject of mined 1-length, is used to identify the relevant features. The two-class base vectors $V_1(\min)$, and $V_2(\min)$,, which represent the deviation values for the full chunk, are identified by carrying out the balanced evaluation of the specific area. Therefore, it is necessary to differentiate between the range of deviation for the two subjects of highest vectors, $V_1(\max)$, along with $V_2(\max)$,. d_{max} and d_{min} are base vector or most extreme vector of departure along the acquired lines. If there is a divergence at that moment, the reasonable qualities are chosen; if not, they are disposed of. The qualities that are selected successfully for a generation of norms are defined in equation (3) as follows:

$$X = \beta_1, \beta_2, \dots\dots \beta_n ; n \le m \qquad (3)$$

Membership function

The input data is transformed by the function for membership (MF) into a connection (or membership degree) number between 0 and 1. The supplied data is converted into a fuzzy value using the triangle membership approach. Equation (4) provides the principle used to examine the membership values.

$$g(x) = \begin{cases} 0, & if \ x \le i, \\ \frac{x-i}{y-i}, & if \ i \le x \le j \\ \frac{k-x}{k-j}, & if \ j \le x \le k \\ 0, & if \ x \ge k \end{cases} \qquad (4)$$

Construction of fuzzy rules with weights

The development of the fuzzy-based clinical decision support system (CDSS) technology begins with two crucial stages: rule weighting and generation. d_{min} and d_{max} are gained for the decision criteria development from the prior stage that is currently in place. The two vectors, which include every property and are compared to the two classes, are automatically used to create the rulers. The decision rulers for each element are obtained from the variance vector. As an illustration, let's say that the d_{max} and d_{min} geometric elements are 8 and 3. The following are the decision rulers that were developed:

✓ The risk is under 50 if β_1 is < 3.
✓ The risk is greater than 50 if β_1 is > 8.
✓ The risk likelihood is either below or over 50 if β_1 ranges from 3 and 8.

Based on the database, we want to determine whether the rulers are satisfied with the patient for each and everyone that is generated. The patient is represented by the letter (H), the rule by the letter (r), and the H value is found by applying the rule ($r_1 \rightarrow r_2$). The rule can be calculated by using the following formula: $W(r_1 \rightarrow r_2) = \frac{H\ (r \rightarrow r_{12})}{N}$,

where N represents all of the patients in the database, and $r_1 \rightarrow r_2$ is the count is the number of patients who meet the criteria.

Determining rules for weighted fuzzy classifier

The fuzzy principles derived from the data illustrated by numerical conceptual values appear to be quite difficult. Managing values of this kind is essential since they are close to human comprehension, and when compared to numerical value principles, guidelines based on these values are typically more understandable and accountable. The fuzzy set assumption handles these values, where fuzzy leads to several improvements of fuzzy rules. The provided software strategy generates fuzzy values from a collection of numeric characteristic rules by empowering the generation of fuzzy modalities. The decision guideline, derived from the past, consists of "if and after". In the if section, the numerical value is mentioned, and the class identity is decided after that section.

The de-fuzzier, fuzzier, fuzzy rules, and fuzzy interference engine made up the fuzzy logic model.

- ✓ Fuzzifer: Fuzzifcation is the initial step of Fuzzifer. Fuzzification is the process of creating fuzzy sets from the gathered input data sets. The fuzzyfication procedure aims to enable comprehension of a fuzzy state in a rule.
- ✓ Base of fuzzy rules: The fuzzy guidelines are essential to every fuzzy framework and are termed as following the fuzziness of the information sources. The fuzzy rules present an IF–THEN standard-based condition. A whole fuzzy norm was created to limit the current yield variable achieves in the standard base.
- ✓ Inference engine: Fuzzy outputs are understood and generated via reasoning based on the collection of rules specified in a fuzzy norms base.
- ✓ De-fuzzifer: The fuzzy set (the sum of the fuzzy set) utilized as an input to the de-fuzzification procedure, and membership serves in the de-fuzzification process. Fuzzy sets that correspond to a crisp output are applied to acquire the single output number.

Pseudocode for proposed GA-FL algorithm

1. Use rough set theory to extract features from input datasets.
2. Feed the gathered characteristics into the fuzzy logic classifier to train the model and generate classification rules.

    (a) Fuzzification: Use equation (4) to turn crisp data into fuzzy data.

    (b) Fuzzy Rule Generation: Use the fuzzy data to produce fuzzy rules.

- The risk is under 50 if β_1 is < 3.
- The risk is greater than 50 if β_1 is > 8.
- The risk likelihood is either below or over 50 if β_1 ranges from 3 and 8.

    (c) Defuzzification: Transform the hazy regulations into clear ones.
3. Maximize the classification rules by applying the hybrid Genetic algorithms to the model constructed in Step 2.

4. Cross-validate the model with the testing data. The model is evaluated using the measures of sensitivity, specificity, and accuracy.

5. Run statistical tests (designate the tests) to confirm the findings.

The pseudocode first uses rough set theory to extract features from input datasets to improve model performance. Fuzzification is then used to handle uncertainty, producing understandable fuzzy rules, while defuzzification produces unambiguous outputs in a fuzzy

logic classifier. In order to improve the classifier's overall performance, a hybrid genetic algorithm is then used to optimize the classification criteria. The model is evaluated for effectiveness using metrics for specificity, sensitivity, and accuracy; its robustness is ensured by cross-validation with test data. Lastly, statistical tests validate the model's reliability and compare it with other methods, confirming the significance of the findings. All the steps add up to a complete methodology that improves classification accuracy and practical application.

4. Results and discussion

a. Evaluation metrics

The parameters listed below are used to evaluate the diagnostic approach for heart disease.

- ✓ The ratio of genuine positive assessments (TP) to total positive assessments (TP+FN) is known as sensitivity.
- ✓ Specificity is calculated by adding (TN+FP) (number of genuine negative assessments) and dividing this by the total count of negative evaluations.
- ✓ The count value of proper assessments divided by the total assessments equals accuracy, equal to (TN + TP)/ (TN + TP + FN + FP).

b. Accuracy

One essential statistic for assessing the effectiveness of classification models, such as the HDD-GA-FL model for heart disease diagnosis, is accuracy. It shows the percentage of instances in the dataset that have been accurately classified out of all cases. The accuracy metric gives an overall assessment of the model's prediction performance. A model with a higher accuracy value performs better because it correctly identifies a more significant percentage of examples.
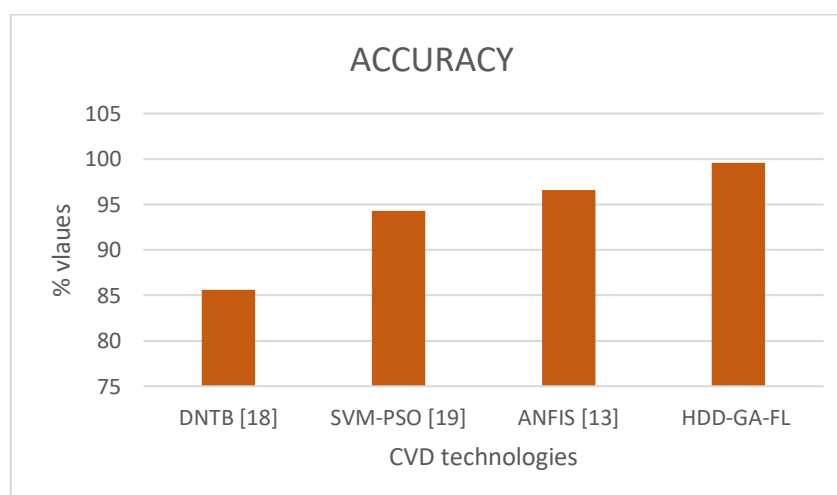

Figure 3. Analysis of the accuracy of the proposed HDD-GA-FL model

The HDD-GA-FL model's accuracy (Figure 3) indicates how well it can diagnose heart disease compared to conventional diagnostic techniques. The model seeks to overcome the difficulties presented by the ambiguity and complexity of coronary artery disease diagnosis by combining the skills of fuzzy logic classifiers and genetic algorithms to effectively categorise instances based on complicated medical data.

The HDD-GA-FL model's efficacy as a diagnostic tool is trusted due to its high accuracy, which indicates that it can consistently distinguish between people with and without

coronary artery disease. Healthcare practitioners need this metric to make educated patient treatment and care decisions, enhance patient outcomes, and better manage heart disease.

c. Sensitivity

A critical performance indicator for binary classification models, such as the HDD-GA-FL model used to diagnose heart disease, is sensitivity, sometimes called the actual positive rate. It measures the model's accuracy in identifying heart disease patients among all positive instances in the dataset. The percentage of accurate positive results the algorithm accurately classifies as positive is known as sensitivity. Stated differently, it assesses the model's capacity to identify coronary artery disease among people with the illness.
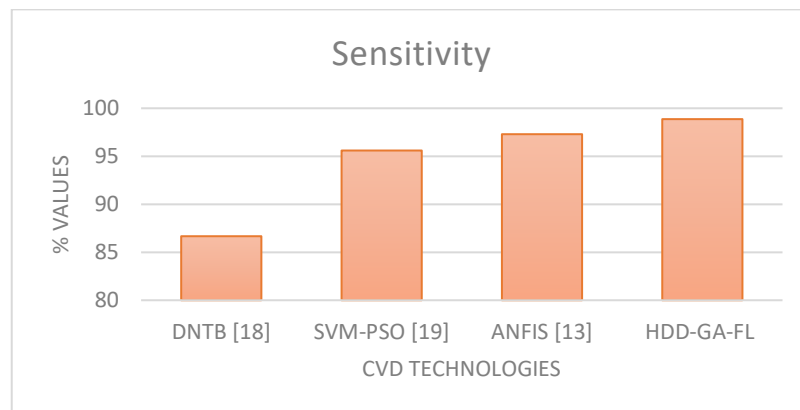

Figure 4. Analysis of sensitivity of proposed HDD-GA-FL

The ability of the model to detect people with heart disease—a critical component of early diagnosis and treatment—is demonstrated by a more excellent sensitivity score illustrated in the Figure. 4. High sensitivity lowers the likelihood of false negatives, thus guaranteeing that patients in need of healthcare receive timely care because the model misses fewer occurrences of heart disease.

A high sensitivity level in the setting of the HDD-GA-FL model would suggest that the model well captures the subtle trends and traits linked to heart disease, making it capable of precisely identifying individuals at risk. This statistic is crucial for medical experts since it directly affects the model's capacity to identify and diagnose heart disease effectively, ultimately enhancing patient outcomes and condition management.

d. Specificity

Another essential performance indicator for binary classification models, such as the HDD-GA-FL model used to diagnose heart disease, is specificity, often called the actual negative rate. It estimates how well the model can identify heart disease-free people among the dataset's negative instances. The percentage of actual adverse events the algorithm correctly classifies as unfavourable is known as specificity. Stated differently, it assesses how well the model can accurately exclude those without coronary artery disease from receiving a positive classification.

A more excellent specificity score means that the algorithm is more effective at detecting people who do not have heart disease, which lowers the possibility of false positive results and pointless treatments. A high specificity reduces the possibility of needless medical

operations or therapies since fewer cases of people without coronary artery disease are mistakenly categorised as positive.
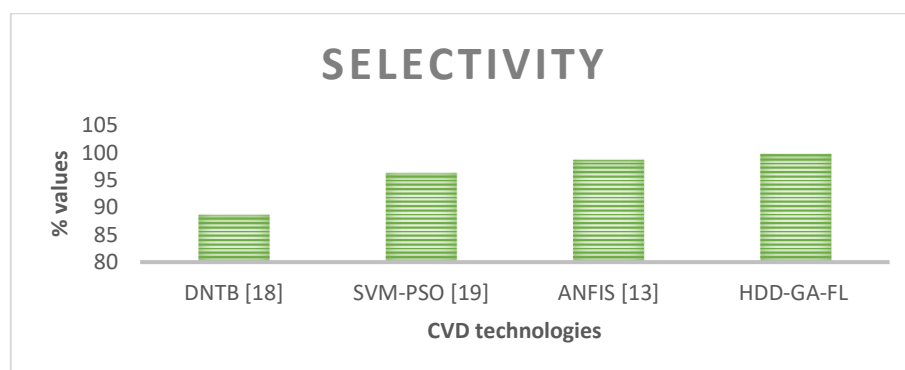

Figure 5: Analysis of selectivity of proposed HDD-GA-FL

A high specificity rating in the setting of the HDD-GA-FL model (as shown in Figure 5) would indicate that the model accurately identifies those who do not need additional testing or treatment, as well as those who have heart disease. This measure is crucial for healthcare workers because it guarantees that resources are used effectively and that interventions are directed toward those who require them. This improves patient care and the management of healthcare resources.

5. Conclusion

Diagnosing heart disease is a crucial area of healthcare since heart disease has a significant impact on death rates globally. To improve Heart Disease Diagnosis (HDD), the study suggests a novel approach called the HDD-GA-FL model, which combines a hybrid genetic algorithm (GA) and fuzzy logic classifier (FL). The proposed hybrid system attempts to address the challenges caused by the complexity and ambiguity associated with cardiac illness diagnosis. When analysing unclear medical data, fuzzy logic classifiers are employed, and genetic algorithms are used for feature selection and optimisation. Combining these two approaches provides a solid basis for accurate and efficient diagnosis. Experiments on a sizable dataset containing cases of cardiovascular disease and other clinical factors are used to assess the efficacy of the hybrid approach. Diagnosis accuracy and reliability have significantly improved compared to traditional methods. GA and FLC work better together than separately when negotiating the complicated feature space in identifying heart sickness. It can pick up on subtle relationships and patterns in the data. The proposed hybrid approach offers a significant promise for practical use in clinical settings, giving physicians a valuable tool for early and accurate diagnosis of heart issues. This approach leverages the combined strengths of genetic algorithms and fuzzy logic classifiers to improve the most recent advancements in cardiac healthcare, ultimately leading to better patient outcomes and cheaper healthcare expenditures. Future research should improve the HDD-GA-FL model's diagnostic skills and adaptability to changing healthcare data by investigating new optimisation strategies and integrating cutting-edge machine learning methods.

REFERENCES

[1] Hameed, Abdul Zubar, et al. "Efficient hybrid algorithm based on genetic with a weighted fuzzy rule for developing a decision support system in predicting heart diseases." *The Journal of Supercomputing* 77 (2021): 10117-10137.

[2] Abdollahi, Jafar, and Babak Nouri-Moghaddam. "A hybrid method for heart disease diagnosis utilising feature selection-based ensemble classifier model generation." *Iran Journal of Computer Science* 5.3 (2022): 229-246.

[3] Kaur, Jagmohan, and Baljit S. Khehra. "Fuzzy logic and hybrid-based approaches for the risk of heart disease detection: a state-of-the-art review." *Journal of The Institution of Engineers (India): Series B* 103.2 (2022): 681-697.

[4] Ashri, Sarria EA, Mostafa M. El-Gayar, and Eman M. El-Daydamony. "HDPF: heart disease prediction framework based on hybrid classifiers and genetic algorithm." *ieee access* 9 (2021): 146797-146809.

[5] Wiharto, Wiharto, et al. "Hybrid Feature Selection Method Based on Genetic Algorithm for the Diagnosis of Coronary Heart Disease." *Journal of information and communication convergence engineering* 20.1 (2022): 31-40.

[6] Eisa, Mohammed M., and Mona H. Alnaggar. "Hybrid rough-genetic classification model for IoT heart disease monitoring system." *Digital Transformation Technology: Proceedings of ITAF 2020*. Springer Singapore, 2022.

[7] Abdollahi, Jafar, and Babak Nouri-Moghaddam. "Feature selection for medical diagnosis: Evaluation for using a hybrid Stacked-Genetic approach in the diagnosis of heart disease." *arXiv preprint arXiv:2103.08175* (2021).

[8] Altae, Aymen A., and Abdolvahab Ehsani Rad. "Diagnosing heart disease by a novel hybrid method: Effective learning approach." *Informatics in Medicine Unlocked* 40 (2023): 101275.

[9] Ramesh, B., and Kuruva Lakshmanna. "A Novel Early Detection and Prevention of Coronary Heart Disease Framework Using Hybrid Deep Learning Model and Neural Fuzzy Inference System." *IEEE Access* 12 (2024): 26683-26695.

[10] Jha, Rahul Kumar, et al. "Personating GA Neural Fuzzy Hybrid System for Computing HD Probability." *International Journal of Advanced Computer Science and Applications* 14.7 (2023).

[11] Baskar, S., et al. "Hybrid fuzzy based spearman rank correlation for cranial nerve palsy detection in MIoT environment." *Health and Technology* 10.1 (2020): 259-270.

[12] Li, Xiaohua, Jusheng Zhang, and Fatemeh Safara. "Improving the accuracy of diabetes diagnosis applications through a hybrid feature selection algorithm." *Neural processing letters* 55.1 (2023): 153-169.

[13] Taylan, Osman, et al. "Early prediction in classification of cardiovascular diseases with machine learning, neuro-fuzzy and statistical methods." *Biology* 12.1 (2023): 117.

[14] Menshawi, Alaa, et al. "A Hybrid Generic Framework for Heart Problem diagnosis based on a machine learning paradigm." *Sensors* 23.3 (2023): 1392.

[15] Ali, Md Liakot, Muhammad Sheikh Sadi, and Md Osman Goni. "Diagnosis of heart diseases: A fuzzy-logic-based approach." *Plos one* 19.2 (2024): e0293112.

[16] Manikandan, G., et al. "Classification models combined with Boruta feature selection for heart disease prediction." *Informatics in Medicine Unlocked* 44 (2024): 101442.

[17] Parveen, Huma, Syed Wajahat Abbas Rizvi, and Raja Sarath Kumar Boddu. "Enhanced Knowledge Based System for Cardiovascular Disease Prediction using Advanced Fuzzy TOPSIS." *International Journal of Intelligent Systems and Applications in Engineering* 12.11s (2024): 570-583.

[18] Alanazi, Sultan Munadi, and Gamal Saad Mohamed Khamis. "Optimizing Machine Learning Classifiers for Enhanced Cardiovascular Disease Prediction." *Engineering, Technology & Applied Science Research* 14.1 (2024): 12911-12917.

[19] Alshraideh, Mohammad, et al. "Enhancing Heart Attack Prediction with Machine Learning: A Study at Jordan University Hospital." *Applied Computational Intelligence and Soft Computing* 2024 (2024).

[20] Kalaivani, B., and A. Ranichitra. "Optimizing Cardiovascular Disease Prediction: Harnessing Random Forest Algorithm with Advanced Feature Selection." (2024).

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