**Performance evaluation and design of the knowledge-based system using advanced fuzzy techniques.**

**Abstract**

The prevalence of cardiovascular disease has risen over the last decade, making it the biggest mortality in the majority of nations. Early detection of heart diseases help the doctors to provide better treatment to the patient. In this research, a knowledge-based hybrid heart disease detection model using advanced fuzzy techniques and Artificial Neural Network (ANN) is proposed. The ANN and advanced fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) techniques are implemented in the proposed methodology for risk prediction of disease and disease classification, respectively. The Analytic Hierarchy Process (AHP) method's attribute weights are helpful for making effective diagnoses of diseases. The proposed model (ANN+ fuzzy TOPSIS) is trained based on different learning percentages such as 60% 70%, 80% and 90% and then tested on every learning percentage also as learning percentage of trained set increases the efficiency of the proposed model is also increased. Various performance measuring metrices are calculated and then compared with other conventional methods to investigate the efficiency of the proposed model. Numerical analysis of the proposed model shows that it performed better than other conventional methods in terms of accuracy (0.99), precision (0.98), specificity (0.978), F-measure (0.981), sensitivity (0.996) and many more.

**Keywords:** Heart disease, advance fuzzy TOPSIS, Data Mining, Information entropy, Knowledge-based system.

1. **Introduction**

There has been a noticeable rise in the prevalence of heart disease, and it has now overtaken all other forms of mortality as the top reason for death for individuals in the majority of nations all over the globe [1]. Many different aspects of cardiovascular disease (CVD) might harm either the anatomy or functionality of the heart [2]. It might be challenging for medical professionals to make a rapid and accurate diagnosis of certain conditions [3-4].

The treatment of cardiac disease is being performed by several systems, many of which depend on methods of soft computing that have been developed [5]. In particular, the combination of multiple different forms of soft computing to construct hybrid models has been examined as a means of producing results that are superior to those produced by a single kind of computational model [6]. In most cases, these models included two distinct states. In the first stage, approaches for selecting features are employed to pick a subset of those characteristics [7]. After that, the produced subset of characteristics is utilized as data for the categorization procedures that are employed in the second state [8]. Figure 1 illustrates a chart showing various types of CVD.

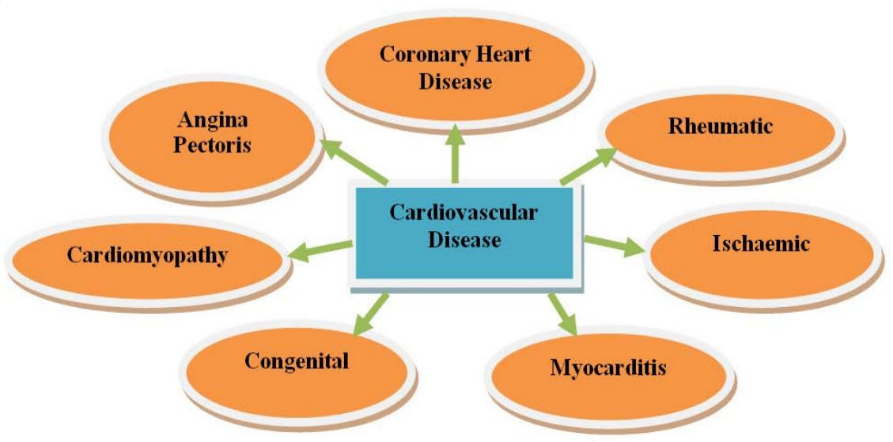


Figure 1 Types of CVD [9]

1. **Fuzzy technique in the healthcare sector**

A set is said to be fuzzy if it permits its components to have varying levels of inclusion in the range [0, 1] [10]. The fuzzy categorization system provides an alternative crisp logic by analyzing data sets according to the individuals' membership in each category [11]. The concept of fuzzy membership is predicated on the idea that a person's participation in a particular group can vary from full membership (100%) to non-participation (0%), and it recognizes the possibility that a dataset might be divided into partial participation in two or more groups [12]. Figure 2 shows an illustration of fuzzy logic.

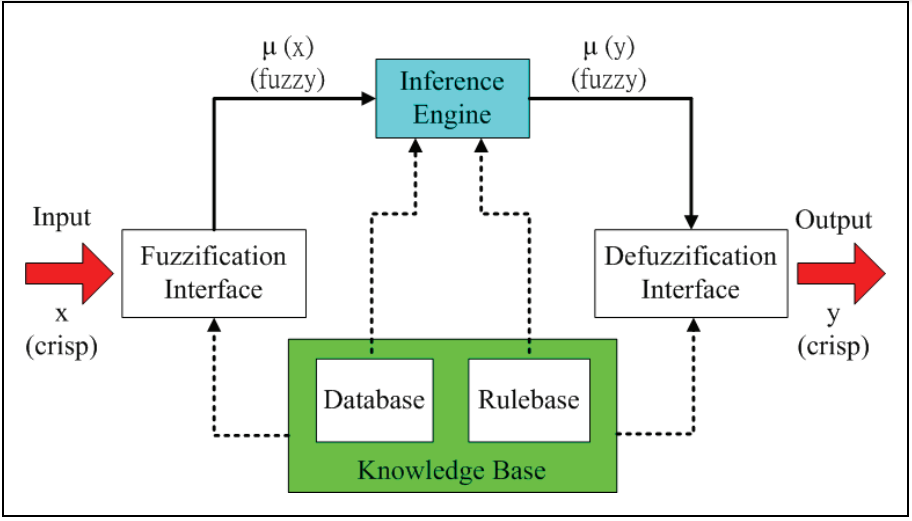


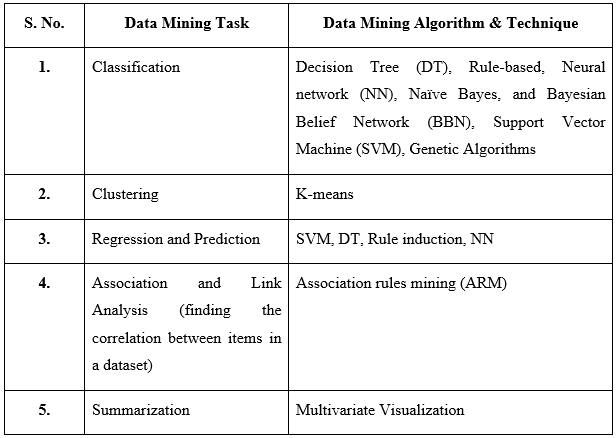
Figure 2 Block diagram of fuzzy logic [13]

To represent the level of membership, fuzzy logic employs truth levels that range from 0.0 to 1.0 [14]. The values of the attributes are changed to fuzzy values. As an example, revenue is projected into the discrete classifications "low, medium, and high," and then fuzzy values are determined for each category. It is possible that more than one fuzzy number would be relevant to a certain fresh sample. Each criterion that is relevant casts a vote about membership in the respective groups. Typically, one would begin by adding up the truth values of each anticipated category [15].

**Data mining technique for heart disease detection**

Data Mining is primarily focused on the evaluation of records, and Data Mining methods and strategies are employed to discover patterns from within a data collection [16]. The primary purpose of data mining is to discover patterns in information in an automated manner, requiring as little work and input from users as possible. Mining data is a strong technique that can be employed to handle decision-making problems [17].

Table 1 Data Mining Tasks and Intelligent Techniques [15]



Current methods of managing and accessing the massive amounts of data generated by healthcare procedures are inadequate [18]. Therefore, the accurate treatment of heart disease requires the application of data mining to retrieve the relevant patterns and information contained inside medical databases [19-20].

1. **Literature review**

This section provides a literature review on the topic of “Performance evaluation and design of the knowledge-based system using advanced fuzzy techniques.”

**Wojcik et al., (2023) [21]** developed a health expert system for determining the severity of coronary artery tumors in individuals who suffer from coronary artery disease using fuzzy sets as the underlying data structure. The use of actual data in testing the intelligent system. In the end, it was found that the level of structural abnormality of the coronary artery in individuals with different kinds of coronary disease was 95%, according to the opinion of specialists.

**Taylan et al., (2023) [22]** discovered that early and correct identification of cardiovascular diseases (CVDs) is of utmost significance to reduce the likelihood of suffering a myocardial attack. As a result, an approach that makes use of the adaptive neuro-fuzzy inference system (ANFIS) technology has been presented as a means of predicting, classifying, and enhancing the diagnostic efficiency of CVDs. According to the findings of the numerical study, the level of accuracy of prediction offered by ANFIS throughout the training phase is 96.56%.

**Kharya et al., (2023) [23]** initiated a novel idea known as the fuzzy-weighted Bayesian belief network (FWBBN), which was used to construct and create a healthcare diagnosis support tool based on the BBN. The fuzzy concept is being used for characteristics to cope with real-life circumstances to eliminate sharp boundary concerns that exist in the medical field. In conclusion, it was determined that FWBBN, in comparison to the traditional Bayesian model, is capable of being applied in a manner that is both highly effective and precisely precise in terms of high efficiency and low time intricacy.

**Seslier and Karakus (2023) [24]** investigated that About 46% of the death of people in the world, excluding communicable diseases and accidents, are because of CVDs. In this analysis various machine learning methods are used to determine heart disease. At last, it concluded that among various classifiers such as Logistics regression, SVM, Naïve bayes, and Random Forest (RF), the SVM technique achieved the best accuracy outcomes at 87.91%.

**Nadakinamani et al., (2022) [25]** evaluated a wide variety of cutting-edge machine learning techniques to develop a CVD forecasting method that is very reliable. To determine which machine learning approach is the most appropriate, the efficiency of the suggested CDPS was measured across several different criteria. The DT approach performed very well, with a maximum accuracy of 100%, when it came to forecasting individuals who would be diagnosed with CVD.

**Cenitta et al., (2022) [26]** recently introduced ischemic heart disease innovative missing value imputation techniques (IHDMIT) using fuzzy-rough sets and their expansions. We evaluate the novel IHDMIT with RF classification against the state-of-the-art methods of expectation maximization, fuzzy C means, and fuzzy roughest. According to the findings, the suggested IHDMIT RF classifier achieves a higher accuracy of 93%.

**Doppala et al., (2022) [27]** discovered that artificial intelligence can transform unprocessed medical data into a useful knowledge base for decision-making and forecasting. The findings of this study offer a robust ensemble model. As a consequence, the suggested model was shown to have an overall success rate of 96.75% on the CVD dataset.

**Yilmaz et al., (2022) [28]** developed three distinct models for classifying coronary heart disease using RF, logistic regression, and SVM methods respectively. Accuracy served as the criterion for determining how well the models performed. At last, it concluded that the RF classifier had the greatest accuracy of 92.9% among all of the classifiers.

Table 2 shows the comparative analysis of the reviewed literature of different authors.

TABLE 2 Comparison of the reviewed literature

|  |  |  |
| --- | --- | --- |
| **Authors Name** | **Technique used** | **Outcomes** |
| **Wojcik et al., (2023) [21]** | Fuzzy set | 95% of patients with different kinds of coronary disease had anatomical coronary artery damage, according to specialists. |
| **Taylan et al., (2023) [22]** | ANFIS | According to a numerical study, ANFIS's training process prediction accuracy is 96.56%. |
| **Kharya et al., (2023) [23]** | FWBBN | Compared to the Bayesian model, FWBBN performs better and takes less time. |
| **Seslier and Karakus (2023) [24]** | SVM | The SVM approach produced the most accurate results, which came in at 87.91% overall. |
| **Nadakinamani et al., (2022) [25]** | DT | The DT approach worked wonderfully in terms of patient prediction for cardiovascular illness, with the greatest accuracy of 100%. |
| **Cenitta et al., (2022) [26]** | IHDMIT | The result demonstrates that the suggested IHDMIT RF classifier provides greater accuracy, 93%. |
| **Doppala et al., (2022) [27]** | Ensemble model | On the dataset for cardiovascular disease, the suggested model yielded results of 96.75%. |
| **Yilmaz et al., (2022) [28]** | RF | In terms of accuracy, the RF classifier has outperformed all others with a score of 92.9%. |

1. **Background Study**

Diverse types of healthcare information have steadily aided in the development of healthcare systems. This data comes from numerous new sources, including computer records, mobile devices, and health monitoring devices. Deep learning can be combined with data fusion techniques to produce more accurate, precise, and reliable forecasts from huge healthcare data sets. This research establishes a four-stage paradigm for predicting risk levels. A novel Ensemble Classifier (EC) with fuzzy technique, Deep Belief Network (DBN), and neural network is presented to anticipate well-being hazards from e- health record of a patient. The fuzzy logic is trained using the returned knowledge sources, while the neural network is trained using the retrieved characteristics. The results of both neural networks and fuzzy logic are sent as inputs to a DBN, which performs the risk estimation for the disorder. To increase the effectiveness of risk factor forecasting, the weight that DBN carries has been adjusted by using a cutting-edge hybrid technique called rain-leveraged dynamic butterfly optimization [29].

1. **Problem formulation**

The traditional techniques for diagnosing CAD essentially make use of the patient's healthcare history, diagnostic test reports, and clinical specialist analysis of the pertinent symptoms. All medical professionals foresee the existence of illness based on information they have gained via their training and experience. Heart disease risk is influenced by a variety of factors, including excessive cholesterol, irregular blood pressure, and inactivity. Because of human error, this might result in a wrong diagnosis and possibly a delay in the prediction of the course and severity of the illness. Such erroneous judgments can prevent timely medical care from being provided or perhaps result in fatalities. As a result, a smart medical decision support system (MDSS) enters the scene and plays a crucial role in healthcare by using patient medical data to forecast illness. In this study, an MDSS is offered as a means of accurately predicting and diagnosing illnesses using a patient's computerized health information.

1. **Research objectives**

* To have a working knowledge of the many types of cardiac disease and the information mining strategies that can be used to prevent them.
* The goal of this project is to create a model for diagnosing cardiac disease based on knowledge-based systems and fuzzy approaches.
* To prove the robustness of the proposed model by comparing it with another conventional model in terms of accuracy and other performance evaluation parameters.

1. **Research Methodology**

The concept of designed architecture is examined in the context of research methodology. artificial neural networks and advanced fuzzy TOPSIS method are used to predict and classify the risk. Further, the proposed model is compared with the conventional technique to prove the robustness and efficiency of the model.

1. **Technique used.**

There are two techniques are used in the proposed methodology. These techniques are given below:

* **Artificial Neural Network (ANN)**

The term ANN refers to a kind of computer model that was developed to imitate the way a human brain operates. Applications involving prediction, categorization, and pattern identification are the primary areas in which ANNs are used. ANN models can understand and recognize similar patterns via training, and then it can recognize the outcome for input data that are presented to the network [30-32].

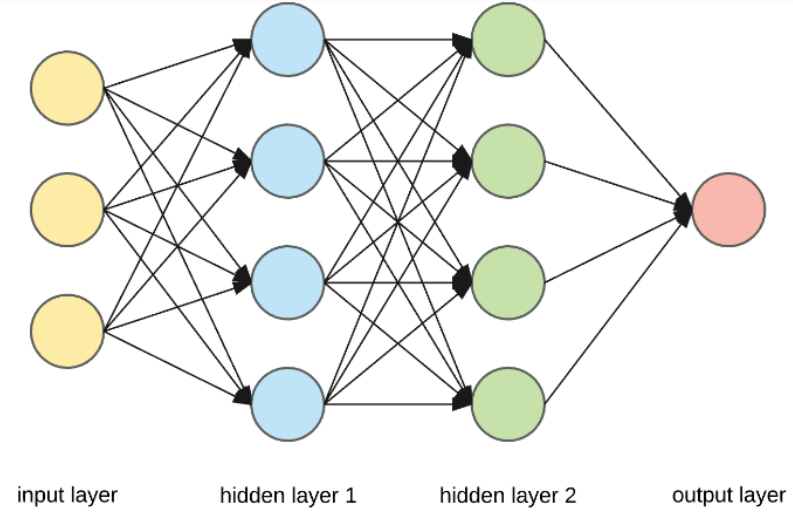


Figure 3 Structure of ANN [31]

* **Fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)**

The fuzzy TOPSIS was first developed by Hwang and Yoon. It is now the method that is the most well-known for handling issues involving the use of many criteria in decision-making. This strategy is predicated on the theory that the selected option ought to have the shortest route to the Positive Ideal Solution, while also having the greatest length to the Negative Ideal Solution [33].The steps taken for fuzzy TOPSIS are as follows:

**Step 1: Ratings should be given both to the conditions and the options.**

In the beginning, it is proceeded on the assumption that the decision-making group consists of K members. The weight of criteria is indicated by , and the fuzzy score of the selection maker on substitute in relation to condition is defined as .

**Step 2: Calculate the accumulated fuzzy ratings for the available options as well as the consolidated fuzzy weights for each criterion.**

The weighted average fuzzy rating of the option with respect to the criteria is calculated as follows: .

(1)

For each criteria , formulae are used to determine the aggregate fuzzy weight .

. (2)

**Step 3 Perform the computation on the normalized fuzzy decision matrix.**

is the formula for the normalised fuzzy decision matrix, where,

The normalized fuzzy decision matrix is, where.

and (3)

and (4)

**Step 4: Perform the computations necessary to produce the weighted normalized fuzzy decision matrix.**

Weighted normalized fuzzy decision matrix is where .

**Step 5: Calculate the Fuzzy Positive Ideal Solution (FPIS) and the Fuzzy Negative Ideal Solution (FNIS). Following are the calculations for FPIS and FNIS:**

where (5)

where (6)

**Step 6: Calculate the length between each option and the FPIS and FNIS.**

The distance between each option and the FPIS and FNIS, respectively, should be given as:

(7)

**Step 7: Calculate the similarity coefficient for each option.**

For each alternative the closeness coefficient is calculated as follows:

(8)

**Step 8: Rank all alternatives.**

The most advantageous option is one with the greatest proximity coefficient [33].

**Proposed Algorithm**

**Start**

1. Read the Electronic health record of the patient as input variable = I

2. Perform data\_ pre-processing (Data cleaning, Tokenization) on → I

3. Weight determination using AHP

4. Select features (I) → using Pearson correlation

5. For → KE, FE; where, KE=knowledge\_ extraction, FE= feature\_ extraction

divide the selected features in group as →g1, g2.

6. Perform→ KE (Ontology-based, Improved semantic similarity) → on g1

7. Perform → FE (statistical features, Information entropy) → using PCA → on g2

8. G1 ← g1+g2

9. Optimize →G1

10. For prediction risk → using→ ANN

11. Classify risk level → using →fuzzy TOPSIS

12. Diagnosis detection

**End**

1. **Proposed methodology.**

Figure 4 shows the flowchart of the proposed methodology. The proposed methodology is further explained step by step in detail.

Input data EHR

Data Pre-processing

(a). Data cleaning

(b) Tokenization

Feature extraction using PCA.

Statistical features

Determining weight using AHP

Knowledge Extraction

Ontology based knowledge extraction.

Improved semantic similarity.

Information entropy

Risk prediction using ANN technique.

Risk level classification using Fuzzy TOPSIS technique.

Diagnosis Detection

Feature selection using Pearson Correlation

Optimization of features

Figure 4 Flowchart of the proposed methodology

**Step 1: Input data EHR**

At the beginning of the process, the Electronic Health Record (EHR) of the patient is taken as the input variable. These health records were further submitted for pre-processing.

**Step 2: Data preprocessing**

In step 2, after getting the EHR as the input variable the preprocessing of data is performed in which the cleaning of data and tokenization is performed so that the noisy data and other inconsistencies can be removed from the raw dataset. As a method of data collection, preprocessing entails standardizing and organizing raw data.

**Step 3: Determining weight using AHP.**

After pre-processing the raw data into organized formed in step 2, in this step, the Analytic Hierarchy Processing (AHP) approach is utilized to generate the attribute weights that are then put to use in the process of initializing the input nodes of a neural network.

**Step 4: Feature selection**

In step 4, after calculating the weight of the organized data using AHP the feature selection using Pearson correlation is performed. With the help of feature extraction, the amount of duplicate data present in a data gathering can be reduced.

**Step 5: Knowledge extraction**

After feature selection, the dataset is divided into two sections. One part is sent for knowledge extraction and the other is sent for feature extraction. In knowledge extraction, the data is extracted based on ontology and then performed semantic similarity. The semantic analysis is done using the following formulae:

(9)

Here, are tokens, are the weights, and are the resources points.

**Step 6: Feature extraction**

In step 6, one portion of data is sent for feature extraction using Principal Component Analysis (PCA). Feature extraction refers to the procedures that can be used to pick and/or aggregate characteristics into features to reduce the amount of information while maintaining the same level of accuracy in the relevant information.

**Step 7: Feature optimization**

In step 7, after performing the knowledge extraction and feature extraction the feature optimization is performed. The data obtained from knowledge and feature extraction is merged and after merging the standard dataset the feature optimization is performed.

**Step 8: Risk prediction using ANN.**

After feature optimization in step 8, the risk prediction using an ANN is performed. Typically, ANNs are used for tasks including prediction, classification, and pattern recognition. It is possible to train an ANN model to recognize a set of similar patterns, and then use that knowledge to predict the network's response to novel inputs.

**Step 9: Risk level classification and decision diagnosis**

After risk prediction using ANN in this last step, the risk classification is completed using the TOPSIS. It is one of the greatest methods for getting an excellent answer out of a set of comparable possibilities. In the selection process, it can also be utilized to optimize the process and eliminate ambiguity and confusion. After risk classification, the decision diagnosis is performed for evaluating the performance of the proposed model.

**Results and discussion**

In this section the results are discussed in detail that are obtained by the implementation of the proposed methodology. The dataset is explained in detail. Performance measured parameters are calculated and finally the proposed model is compared with the conventional technique.

**Dataset**

The dataset that is used in this research is easily available on the website of Kaggle. It is a collection of electronic health records of the patients. The information such as age, gender, chest pain type, blood pressure , ECG., etc. is stored in the records. The dataset is split into two parts. 70 % data is used to train the model and the remaining 30 % is used to test the model. The code implementation is done using python language. Table 3 describe the dataset in detail.

Table 3 Dataset report

Table

Description automatically generated with medium confidence

**Result 1:**

In this analysis the sensitivity of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 4 and figure 5 that as the learning percentage of the dataset is increases the sensitivity of the proposed model is also increased. Sensitivity of the proposed model can be calculated by the following formulae:

(10)

Table 4 Sensitivity at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **Sensitivity** | 0.992 | 0.994 | 0.995 | 0.996 |

Figure 5 Graph showing Sensitivity of proposed model.

**Result 2:**

In this analysis the specificity of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 5 and figure 6 that as the learning percentage of the dataset is increases the specificity of the proposed model is also increased. Specificity of the proposed model can be calculated by the following formulae:

(11)

Table 5 Specificity at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **Specificity** | 0.975 | 0.976 | 0.978 | 0.978 |

Figure 6 Graph showing Sensitivity of proposed model.

**Result 3:**

In this analysis the accuracy of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 6 and figure 7 that as the learning percentage of the dataset is increases the accuracy of the proposed model is also increased. Accuracy of the proposed model can be calculated by the following formulae:

(12)

Table 6 accuracy at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **Accuracy** | 0.981 | 0.983 | 0.986 | 0.99 |

Figure 7 Graph showing Accuracy of proposed model.

**Result 4:**

In this analysis the precision of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 7 and figure 8 that as the learning percentage of the dataset is increases the precision of the proposed model is also increased. Precision of the proposed model can be calculated by the following formulae:

(13)

Table 7 Precision at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **Precision** | 0.962 | 0.970 | 0.974 | 0.98 |

Figure 8 Graph showing Precision of proposed model.

**Result 5**

In this analysis the F-measure of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 8 and figure 9 that as the learning percentage of the dataset is increases the sensitivity of the proposed model is also increased. F-measure of the proposed model can be calculated by the following formulae:

(14)

Table 8 F-measure at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **F-measure** | 0.912 | 0.981 | 0.982 | 0.981 |

Figure 9 Graph showing Sensitivity of proposed model.

**Result 6**

In this analysis the Mathews Correlation Coefficient (MCC) of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 9 and figure 10 that as the learning percentage of the dataset is increases the MCC of the proposed model is also increased.

Table 9 MCC at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **MCC** | 0.96 | 0.963 | 0.971 | 0.98 |

Figure 10 Graph showing MCC of proposed model.

**Result 7:**

In this analysis the Negative Predictive Value (NPV) of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 10 and figure 11 that as the learning percentage of the dataset is increases the NPV of the proposed model is also increased.

Table 10 NPV at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **NPV** | 0.994 | 0.995 | 0.998 | 1.0 |

Figure 11 Graph showing NPV of proposed model.

**Result 8**

In this analysis the False Positive Rate (FPR) of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 11 and figure 12 that as the learning percentage of the dataset is increases the FPR of the proposed model is decreased.

Table 11 FPR at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **FPR** | 0.024 | 0.020 | 0.016 | 0.015 |

Figure 12 Graph showing FPR of proposed model.

**Result 9**

In this analysis the False Negative Rate (FNR) of the proposed model is calculated on different learning percentage such as 60%, 70%, 80% and 90%. It is clearly seen in table 12 and figure 13 that as the learning percentage of the dataset is increases the FNR of the proposed model is also increased.

Table 12 FNR at different learning percentage

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **Learning percentage** | | | |
| **60%** | **70%** | **80%** | **90%** |
| **FNR** | 0.008 | 0.006 | 0.004 | 0.003 |

Figure 13 Graph showing FNR of proposed model.

* **Comparative analysis**

In this section, the proposed model is compared with other conventional methods such as Neural network (NN), SVM and DBN. It is compared on the basis of positive metrics parameters such as sensitivity, specificity, accuracy, and precision. Figure 14 (a) shows the comparison of conventional technique with the proposed model on the basis of sensitivity, and it is clearly seen the sensitivity of the proposed model is higher among all the methods. Figure 14 (b) shows the comparison of conventional technique with the proposed model on the basis of specificity, and it is clearly seen the specificity of the proposed model is higher among all the methods. Figure 14 (c) shows the comparison of conventional technique with the proposed model on the basis of accuracy, and it is clearly seen that the accuracy of the proposed model is higher among all the methods. Figure 14 (d) shows the comparison of conventional technique with the proposed model on the basis of precision, and it is clearly seen the precision of the proposed model is higher among all the methods. Table 13 shows the overall comparison of the proposed model with other conventional techniques in terms of sensitivity, specificity, accuracy, and precision.

Table 13 Comparison table

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Sensitivity** | | | | **Specificity** | | | | | **Accuracy** | | | | **Precision** | | | |
| **60** | **70** | **80** | **90** | **60** | **70** | **80** | | **90** | **60** | **70** | **80** | **90** | **60** | **70** | **80** | **90** |
| **NN** | 0.581 | 0.593 | 0.639 | 0.580 | 0.895 | 0.898 | 0.909 | | 0.895 | 0.832 | 0.837 | 0.855 | 0.832 | 0.581 | 0.593 | 0.639 | 0.580 |
| **SVM** | 0.726 | 0.740 | 0.772 | 0.739 | 0.725 | 0.735 | 0.759 | | 0.735 | 0.744 | 0.757 | 0.786 | 0.757 | 0.491 | 0.516 | 0.573 | 0.516 |
| **DBN** | 0.991 | 0.992 | 0.994 | 0.992 | 0.971 | 0.973 | 0.976 | | 0.973 | 0.979 | 0.981 | 0.984 | 0.981 | 0.960 | 0.963 | 0.971 | 0.963 |
| **ANN+TOPSIS** | 0.992 | 0.994 | 0.995 | 0.996 | 0.975 | 0.976 | 0.978 | | 0.978 | 0.981 | 0.983 | 0.986 | 0.99 | 0.962 | 0.97 | 0.974 | 0.98 |
|  | | | | | | | |  | | | | | | | | | |
| (a) | | | | | | | | (b) | | | | | | | | | |
|  | | | | | | | |  | | | | | | | | | |
| (c) | | | | | | | | (d) | | | | | | | | | |

Figure 14 Comparison of the proposed work's performance to that of similar current schemes in terms of (a) Sensitivity (b) specificity (c) accuracy (d) precision

In figure 15 (a) and (b), a comparison of the proposed model with other conventional technique is shown on the basis of FPR and FNR. Figure 15 (a) shows that the overall FPR of the proposed model is quite low as compared to other technique and figure 15 (b) shows that the overall FNR of the proposed model is quite low as compared to other techniques. Table 14 shows the overall comparison of the proposed model with other conventional techniques in terms of FPR and FNR.

Table 14 Comparison table

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **FPR** | | | | | **FNR** | | | |
| 60 | **70** | **80** | **90** | | **60** | **70** | **80** | **90** |
| **NN** | 0.104 | 0.101 | 0.090 | 0.104 | | 0.418 | 0.406 | 0.360 | 0.419 |
| **SVM** | 0.274 | 0.264 | 0.240 | 0.264 | | 0.273 | 0.259 | 0.227 | 0.260 |
| **DBN** | 0.028 | 0.026 | 0.023 | 0.026 | | 0.008 | 0.007 | 0.005 | 0.007 |
| **ANN+**  **TOPSIS** | 0.024 | 0.02 | 0.016 | 0.015 | | 0.008 | 0.006 | 0.004 | 0.003 |
|  | | | | |  | | | | |
| (a) | | | | | (b) | | | | |

Figure 15 Comparison of the proposed work's performance to that of similar current schemes in terms of (a) FPR (b) FNR

In figure 16 (a), (b), and (c) a comparison of the proposed model with other conventional technique is shown on the basis of NPV, F-measure and MCC, respectively. Figure 16 (a) shows that the overall NPV of the proposed model is quite low as compared to other technique. Figure 16 (b) shows that the overall F-measure of the proposed model is quite low as compared to other technique Figure (c) shows that the overall MCC of the proposed model is quite low as compared to other technique. Table 15 shows the overall comparison of the proposed model with other conventional techniques in terms of F-measure, MCC, and NPV.

TABLE 15 Comparison table

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | | **F-measure** | | | | **MCC** | | | | | **NPV** | | | |
| **60** | **70** | **80** | **90** | **60** | **70** | | **80** | **90** | **60** | **70** | **80** | **90** |
| **NN** | | 0.581 | 0.593 | 0.639 | 0.580 | 0.477 | 0.491 | | 0.549 | 0.475 | 0.895 | 0.898 | 0.909 | 0.895 |
| **SVM** | | 0.571 | 0.594 | 0.647 | 0.594 | 0.432 | 0.458 | | 0.520 | 0.458 | 0.922 | 0.926 | 0.935 | 0.926 |
| **DBN** | | 0.975 | 0.978 | 0.682 | 0.977 | 0.959 | 0.962 | | 0.968 | 0.962 | 0.994 | 0.994 | 0.995 | 0.994 |
| **ANN+**  **TOPSIS** | | 0.912 | 0.981 | 0.982 | 0.981 | 0.96 | 0.963 | | 0.971 | 0.98 | 0.994 | 0.995 | 0.998 | 1 |
|  | | | | | | |  | | | | | | |
| (a) | | | | | | | (b) | | | | | | |
|  | | | | | | | | | | | | | |
| (c) | | | | | | | | | | | | | |

Figure 16 Comparison of the proposed work's performance to that of similar current schemes in terms of (a) NPV, (b) F-measure, and (c) MCC

**Conclusion**

In most nations, heart disease has surpassed all others as the top cause of mortality among adults during the last decade. Heart disease is a common health problem, and early identification helps physicians treat these patients more effectively. The proposed model achieved encouraging outcomes in predicting and classifying CAD risk. The AHP method's attribute weights can help in making sound decisions about illness diagnosis. By reducing the need for invasive biopsies and clinical procedures, the system protects users from the risks associated with CAD diagnosis. The system provides a second perspective to the doctor in order to confirm the diagnosis of sickness. According to numerical analysis, the suggested model outperformed standard techniques in many respects, including accuracy (0.99), precision (0.98), specificity (0.978), F-measure (0.981), sensitivity (0.996), and many more. In future work, the researcher will carry on their investigation into and development of effective heuristic methods for attribute reduction strategies, with the goal of managing enormous quantities of features and big numbers of records.