Journal of Information Technology Management

Print ISSN: 2008-5893

Online ISSN: 2423-5059

An Accurate Prediction Framework for Cardiovascular Disease Using Convolutional Neural Networks

Krishna Lava Kumar Gopu *

*Corresponding author, Computer Science and Engineering, Geethanjali College of Engineering and Technology, Hyderabad, Telangana, India. E-mail: lavakumar.gopu@gmail.com

Suthendran Kannan

Information Technology, Kalasalingam Academy of Research and Education, Krishnankoil, Srivilliputhur, Tamil Nadu, India. E-mail: k.suthendran@klu.ac.in

Abstract

Cardiovascular-Diseases (CVD) are a principal cause of death worldwide. According to the World-Health-Organization (WHO), cardiovascular illnesses kill 20 million people annually. Predictions of heart-disease can save lives or take them, depending on how precise they are. The virus has rendered conventional methods of disease anticipation ineffective. Therefore, a unified system for accurate illness prediction is required. The study of disease diagnosis and identification has reached new heights thanks to artificial intelligence. With the right kind of training and testing, deep learning has quickly become one of the most cutting-edge, reliable, and sustaining technologies in the field of medicine. Using the University of California Irvine (UCI) machine-learning (ML) heart disease dataset, we propose a Convolutional-Neural-Network (CNN) for early disease prediction. There are 14 primary characteristics of the dataset that are being analyzed here. Accuracy and confusion matrix are utilized to verify several encouraging outcomes. Irrelevant features in the dataset are eliminated utilizing Isolation Forest, and the data is also standardized to enhance accuracy. Accuracy of 98% was achieved by employing a deep learning technique.

Keywords: Deep-Learning, CNN, Heart-Disease, Prediction, Cardiovascular Disease, Accuracy

Journal of Information Technology Management, 2024, Vol. 16, Issue 1, pp. 27-43
Published by University of Tehran, Faculty of Management
doi: https://doi.org/ 10.22059/jitm.2024.96373
Article Type: Research Paper

Received: December 17, 2023
Received in revised form: January 06, 2024
Accepted: February 02, 2024
Published online: February 29, 2024

Introduction

According to a recent study by the WHO, cardiovascular disease kills 20 million people worldwide every year. It's not shocking that prediction puts it at 85 million by 2030. Tobacco use, unhealthy food choices, lack of exercise, and excessive alcohol consumption are all risk factors for cardiovascular-disease (CVD). Heart attacks are common in people with cardiovascular disease (Manimurugan et al., 2022). As can be seen below, early diagnosis of the disease is essential, as is instruction on the use of temporary drugs. In the long run, CVD is resolved when fatty stores in the ducts and blood groups are generated. Injuries to the brain, eyes, heart, or kidneys are additional possible causes. Maintaining a healthy lifestyle can help diminish the risk of CVD, the foremost cause of demise and injury in the United Kingdom (Faieq et al., 2022). Clots that block blood surge to the brain or heart are the leading cause of cardiac arrest and stroke, both of which can be precipitated by traumatic events. The most well-known motivation for this is the desire to foster the most insular greasy joints possible (Budholiya et al., 2022).

These models have been used in the healthcare sector for more specialized purposes, such as disease prediction and detection, since the 1970s. For instance, MYCIN was developed at Stanford University to detect bacterial blood-borne illnesses (Ansarullah et al., 2002). Recent research using IBM's Watson has concentrated on the precise prescription of medication, in particular the accurate diagnosis of cancer and its treatment. Researchers were aided in their efforts to create new software using Google's Tensor Flow. In order to manage and enhance patients' health from a variety of chronic health issues, medical representations and service providers utilize their irrefutable experience to create individualized programs for each patient. Consistent clinic visits, weight loss, and an overall healthier lifestyle are all recommendations that patients should follow regardless of their specific health situation. But serious issues arise when a patient does not adhere to treatment (Ahmed et al., 2022). One of the most pressing concerns among researchers is the ability to foresee the onset of sickness in the human body. Not even medical professionals are good at foreseeing outbreaks of the disease. However, they require a dependable resource for disease prognosis. Some of the algorithms can be used; however the performance of the system as a whole might use some enhancements. Therefore, there is a vast opportunity for study to aid doctors in the prediction of CVD disease in humans (El-Hasnony et al., 2022).

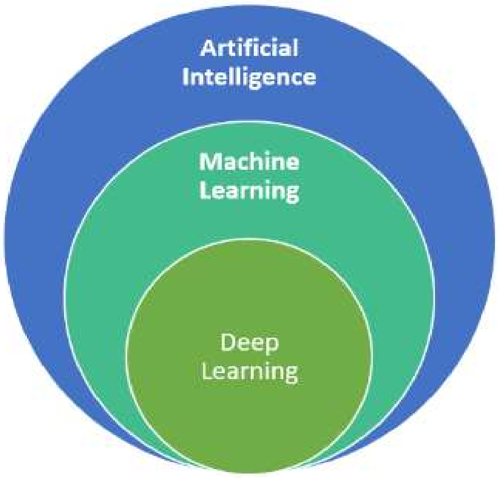

Figure 1. Progression of deep learning

Machine learning is a more narrow type of computational mechanism that uses an algebraic design and instruction data to learn how to produce predictions, while Artificial-Intelligence (AI) applies to various needs of tools that perform activities that are characteristic of human intelligence, such as assuming final thoughts coming from deduction or logical reasoning (Riyaz et al., 2022). ML, as opposed to explicitly calculating results from a set of predefined regulations, learns criteria from examples and thus has the potential to perform better at a task like discovering and classifying information through being exposed to an additional instance. As seen in Figure 1, this is exactly why one of the most cutting-edge ML techniques, deep learning (DL), is being used.

Research into the automatic elimination of complex information portrayals at high levels of absorption has uncovered some promising opportunities, one of which is the use of deep learning models. Such protocols produce a hierarchical structure for knowledge and embodiment in which higher-level parts are determined with reference to lower ones (Ashreetha et al., 2022). Deep Learning protocols' structured learning style is actually motivated by expert system simulations of the layered learning and split accepting procedure of the major sensorial regions of the human brain, which directly draws out components and absorptions from the rooting files. When dealing with vast amounts of unsupervised data, deep learning techniques can be quite helpful, as they often discover record representations in a greedy layer-wise fashion (Lokesh et al., 2022).

Early disease detection is a major challenge in the medical industry. The available data made it impossible to make accurate predictions or draw meaningful conclusions. To this end, we require a decision-making mechanism that can aid medical professionals in reaching an accurate diagnosis of any illness. Heart-based disorders are one example of such perilous conditions (Nancy et al., 2022). Heart-related illnesses kill a staggering number of individuals every year-in the tens of millions. Their everyday routine or the current state of affairs could be to blame. Numerous prediction models have been developed using a variety of datasets, so clearly this is an active topic of study. However, as time goes on, more and more study data and hospital patient records become available. Patients' medical histories can be accessed through numerous publicly available channels, and studies can be undertaken to determine which computer technologies would be most effective in making an accurate diagnosis and preventing the spread of this potentially fatal disease. It is common knowledge at this point that machine learning and AI have substantial practical applications in healthcare. The disease can be classified or predicted using various ML & DL strategies. ML models simplify the process of analyzing genetic material in its entirety. Medical records may be converted and studied more extensively, and models can be taught to make pandemic predictions (Shanmugaraja et al., 2023).

Several publications and studies have shown that AI can do better than humans at healthcare-related tasks. Diseases of varying types and severities are being categorized and identified with the aid of ML &DL strategies (Chitra et al., 2022). ML as a statistical method for modeling data and in order to unlock improved accuracy, training the model using massive types of data is required. CNN, one of the deep learning algorithms, outperforms conventional methods in disease diagnosis. This CNN-based model processes massive amounts of information. CNN's main benefit is that all of the work preprocessing, feature extraction and prediction is handled automatically by the network. Raw data is accepted by the system.

The idea here is to scrutinize how well various ML strategies can foretell cardiac-issues. The CNN model is utilized to construct these forecasting tools and k-modes grouping to preprocess the dataset and scale it to amplify the models' meeting. Python was used for all the computation, preprocessing, and visualization on Google-Colab. ML strategies have been shown to anticipate cardiac disease with up to 94% accuracy (Song et al., 2015). However, these studies frequently employ inadequate sample sizes, thus their findings may not apply to the population at large. Using a broader and more diverse dataset is one way we hope to overcome this restriction in our research and make our findings more applicable to a wider audience.

The following are some of the key contributions that the suggested approach offers:

- Using a cutting-edge DL strategy called CNN, the suggested framework predicts the likelihood that a patient has cardiovascular disease.
- To the best of our knowledge, this is the first time a DL strategy has been used to predict a CVD in the medical profession with only 14 variables.
- We did a decent job of preparing the heart-disease dataset and comparing it to the outcomes from state-of-the-art technologies.

Here is how the rest of the manuscript is laid out: The current literature is described in Sec.2. In Sec.3, we detail the features and benefits of the proposed methods. The experimental findings and analyses are illustrated in Sec.4. In Sec.5, we draw the final conclusions.

Literature Review

As an example, medical images can be used to forecast CVD disorders, and deep learning can be used to classify these images. Predicting cardiovascular imaging wants a number of components, the most essential of which are medical imaging and deep learning algorithms. As a result, it has an effect on how those factors evolve. This section details the numerous proposals made by researchers for predicting cardiovascular disease. The rate of cardiovascular failure with pulse transition was researched extensively by (Almustafa et al., 2020) using time analysis, ML, and CNN models for investigation. In a test of cardiovascular recognition using several classification methods, the support vector machine came out on top. It was hypothesized that ML models may predict the rate of CVD events in patients with cruel dilated cardiomyopathy (DCM) over the course of a year. Information Gain (IG) picked out 32 highlights from the clinical data that were most related to cardiovascular events, which was the ML algorithm's contribution. Human cardiovascular infections treated with antibiotics were investigated. Base of Fisiologiab Clinica has utilized and considered two ML approaches; the other was a dataset backed by the American Diabetes Association, the American Stomach Association, and the Vault of Kidney Diseases.

The use of artificial-neural-networks (ANN) for cardiac infections was anticipated by KM (Ahmed et al., 2017). Machine learning and pattern matching techniques will be used to find a treatment for heart disease. The Continuous-Cardiovascular Disruption Detection of Heart Rhythm was predicted using a set of ML classifiers. Prediction methods like these use strategies including sampling, segmentation, and function selection. This study provides an approach to risk prediction for cardiovascular disease using computational machine learning. In order to choose and put into action the ML simulation pipelines, researchers used an ML-built model derived from auto-prognosis and an algorithmic technique. The ML technique presented by (LK et al., 2021) is another effective tool for diagnosing coronary artery disease. In this method, we present a new technique for advancing, which we dub the N2 Genetic Optimizer Agent. These kinds of tests are quite violent and are comparable to the most reliable findings in the field.

The Continuous Arrhythmia-heartbeat recognition framework was supplied by (Huang et al., 2023). The treatment employs the parallel delta modulations and the linear support-vector-machine (SVM). Improved fluorescence imaging with photonic crystals investigated the use of machine learning in immunoassay biomarker testing for cardiovascular disease. Principle component analysis, partial least squares regression algorithms, and conventional data mining strategies are used (Alizadeh et al., 2022) examined machine learning's application to coronary artery disease; the method breaks down the fundamental steps by testing datasets, analyzing the weights, measuring the implementation, and using ML. Classifier systems trained by machine learning were employed. When compared to other classification models, the random forest classifier performed best in a hepatitis prediction study.

Authors (Saqlain et al., 2019) used linear regression and random-forest, with accuracies of 80% &60% respectively, to predict painful acknowledgement of failing to meet with varying specifics of patients with cardiovascular health concerns. (Gavhane et al., 2018) presented the repository of ML formulae and tools for CVD investigation, anticipation and provided a comparative evaluation of several algorithms in the form of a research paper. Using the UCI heart condition dataset for presenting their work predicting CVD with K-means jumble and predicting a hybrid recommendation. They introduced a multilayer perceptron ANN model with a hidden layer of 18 neurons. When applied to a testing population, this design has achieving improved accuracy than conventional mdels. Using the UCI dataset for heart disease prediction via probabilistic principal component analysis (PPCA), introduced a hybrid method that takes clinical assessment findings as input and extracts an abridged dimensional property.

For cardiac issue prediction, (Muser et al., 2019) used a calculated backbreeding multilayer perceptron (MLP). Predicting cardiovascular disease using neural network classification was a method proposed. They also worked on showing the individual's risk levels utilizing models like K-nearest neighbor, decision tree, &naive byes. They introduced anticipation design that included many elements, each of which could be combined in a number of ways, and a handful of well-established clustering methods. Using a hybrid machine learning approach, the authors improved performance to an accuracy of 84.42%. Using the CVD dataset, (Doppala et al., 2022) found that the expected eminence option strategy had implemented a precision of 81.23% with SMOTE and the XGBoost classifier for a variety of clinical functions. A novel feature selection algorithm was proposed by the authors. Meanwhile, data discretization is used to deal with the uncertainty in coronary artery disease (CAD) prediction.

(Bharti et al., 2021) built an ensemble design that, when combined with current AI techniques, is substantially more accurate at identifying cardiac-based disorders. Using data from actual health care facilities, they implemented a forecasting design. Using both organized and unstructured data on individuals, the authors propose a CNN strategy as a disease prediction algorithm. The established model yields an accuracy of between 85% and 88%. To improve the COVID-19 setup, (Baskar et al., 2023) developed a CNN model optimization. The design is developed and used to categorize the infected population. With a 10-fold relevance cross-validation, predicted a strategy that outperforms all of the base classifiers in the dataset. In terms of accuracy (93.55 percent), our discovery approach has outperformed the already available versions which kept the old classifier sets and private classifiers. On a large upper body CT dataset, the suggested ensemble design was evaluated against fifteen low-cost alternatives. The theoretical findings show that the suggested set version has better accuracy, F-measure, AUC, sensitivity, and specificity than the existing designs by 1.27 percentage points, 1.32 percentage points, 1.28 percentage points, and 1.83 percentage points, correspondingly.

(Reddy et al., 2021) are developed a scheme that feeds data to an android app's analytics. After the model has been evaluated, a pre-trained tool is used to gain familiarity with it; this machine was trained using the same dataset that was used during deployment to Firebase. Finally, LR is implemented in disease diagnosis. In a study comparing 6 ML models, they found that the neural network model had the highest accuracy, at 93%. Finally, a novel SVM-based architecture for cardiac disease prediction was described. With an accuracy of 96.23 percent, the suggested design outperforms the traditional by a wide margin. In their discussion of disease detection, (Kumar et al., 2023) note that logistic regression was used to make predictions in their research. Extensive experimental results reveal that, for the COVID-19 dataset, the proposed version outperforms the competitive equipment finding versions in terms of precision (1.4765%) and F-measure (1.2782), respectively. For the diabetic mellitus dataset, the advised version outperforms the inexpensive device discovering variants in terms of precision (1.8274%) and F-measure (1.7264).

Methodology

Computerized heart disease prediction is the focus of this research, as it can help both doctors and patients plan preventative care. In this article, we show how we accomplished this by applying several machine learning techniques to a dataset. We want to improve the process by standardizing the data, removing extraneous information, and adding new elements like mean arterial pressure (MAP) and body mass index (BMI) (Muthappa et al., 2023). After that, we'll use k-modes clustering to create a gender-specific subset of the dataset. At last, we'll use the cleansed data to train the model. Figure 2 shows that the revised process will lead to more precise results and better model performance. It is now more important than ever to collect, store, search, and share the vast volumes of data being produced by a wide range of sources, yet this data is also notoriously difficult to understand and analyze. Researchers have been exploring for ways to increase the model's accuracy and deliver improved disease prediction results in light of the massive amount of data and the rising cost of diagnosis (Latha et al., 2019).

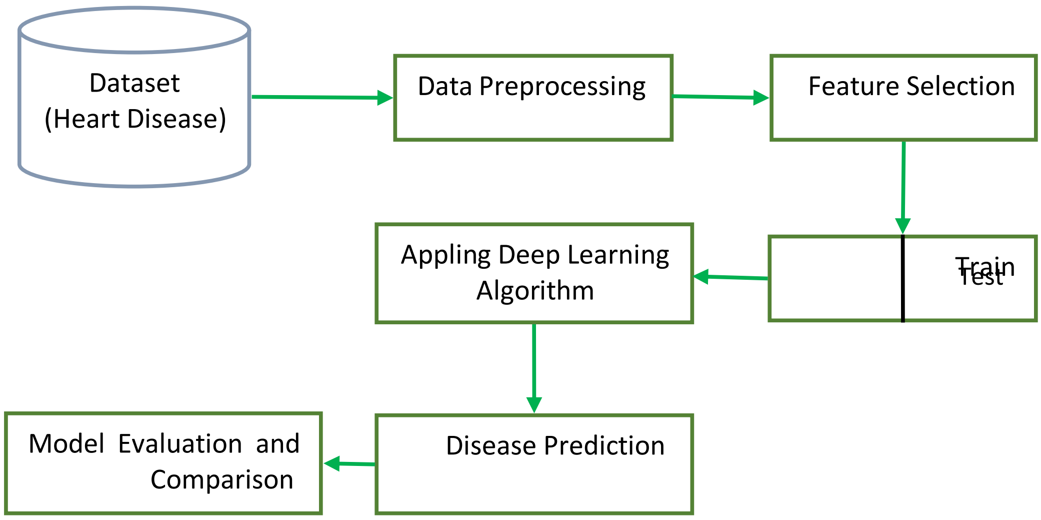

Figure 2. Concept diagram of proposed Model

Data Preprocessing

For this research, we used the Cleveland repository's UCI dataset on cardiovascular illness. From the UCI repository, which has 303 records with fourteen characteristics each, we were able to compile this catalog. For this data collection, we utilized the R program for statistical analysis. Attribute target is the heart disease class attribute this prediction algorithm uses. Data preprocessing is a technique used in data mining that involves transforming unstructured data into a more human-friendly format. The data we get from the real world is sometimes inaccurate, incomplete, and/or lacking in relevant behaviors and/or trends. One tried and true method of dealing with this issue is through preprocessing data (Sampath S et al., 2023). The term "data preprocessing" refers to the steps taken before any actual processing can begin on the data. During data preparation, missing variables are populated with mode values determined by the specific data collection. Data Origination Second, outliers are not removed from the dataset (those with heart disease whose variables have extreme values).

Feature Selection

In order to eliminate unnecessary features and boost performance in order to construct a superior classification model, feature selection is a crucial step in the data preprocessing process. The dataset undergoes feature selection so that a subset of relevant characteristics may be chosen for model construction; this is done so that model accuracy can be enhanced. Feature selection is a powerful method for reducing the dimensionality of data in data mining. Finding the most important risk variables associated with an illness is essential in medical diagnosis. Eliminating unnecessary information from the sickness dataset is sped up and improved by identifying relevant features (Suneel et al., 2024).

Convolutional Neural Networks (CNN)

Predicting whether or not a patient has heart-disease is an example of a binary classification problem. In the context of supervised learning, the NN has been shown to perform well as a classifier (Chaurasia et al., 2013). Recent studies have shown that neural networks tuned to specific applications, such as those with several hidden layers, can significantly outperform their generic counterparts. Image-processing, speech-processing, and time-series anticipation are all domains where neural networks have been successfully applied (Mamatha et al., 2023). Intense training and tweaking of various deep learning architectures have integrated relatively larger datasets. The input data is transformed over the hidden-layers, and the error is estimated at the output-layer in an artificial neural network. The output layer's backpropagated error is then used in a gradient descent algorithm's iterative update of subsequent layers' weights. Reducing overfitting, timing the training-process, building the layers nonlinear, visualizing the hidden-layers, and other adjustments have all been offered as ways to improve the gradient descent algorithm based on various experiments and assessments. Despite significant success in its applications, little is known about how deep neural networks actually work. Millions of parameters in a deep architecture make it easy for the training networks to be overfit. When there aren't enough instances, the issue becomes much more complicated.

Many algorithms have been offered as potential solutions to this problem. One of these common practices is data augmentation, which uses existing instances to artificially create new, smaller datasets. Although this method produces more believable instances overall, it lacks credibility when applied to biological applications like clinical datasets. In some cases, the expanded CVD phenotypic parameters, including platelet count, may not fall within the patient's normal range of values. The discrepancy between the theoretical underpinnings of platelet count readings and the rules of statistical-generation is the root cause of this situation. Unreliable and erroneous categorization results from insufficient training on small or skewed datasets. When contrasted to other applications and the consequences of making an incorrect prediction in medical research are much more severe. An untreated CVD patient may receive the incorrect therapeutic medicine due to an ineffective prediction strategy. One of the primary goals of this study is to enhance the correctness of the classification, i.e., to better forecast the presence or absence of CVD in a subject, because the exactness of an anticipation approach is crucial in medical-application. To address these issues, we advocate for the use of a shallow CNN. As can be seen in figure3, the convolution layers of the aforementioned shallow CNN are sandwiched between two fully linked layers.

The proposed CNN model makes use of dropout layers anyway. To further lessen the likelihood of overfitting, we employed this training regimen. After training the model with a 1: N weight-ratio for a long enough time, the weight ratio is increased while the number of epochs is gradually decreased. The primary goal of the training-layer is to generate a trained model for cardiovascular disease prediction. The provided training data's labeled outputs can be used to train the learning model. Only 30% of the rows are used for actual training, while the remaining 70% are used for actual testing and validation. There are three distinct operations that make up the training layer: the leading and cleansing operation, the model training operation, and the storing operation for the trained model. Each stage is briefly explained below. The heart disease prediction-layer centers on testing data sets for accuracy. The prediction and testing phase involves providing the training model with the remaining70% of the information (without output labels). The primary goal of the training-layer is to generate a trained aproach for cardiovascular disease prediction.

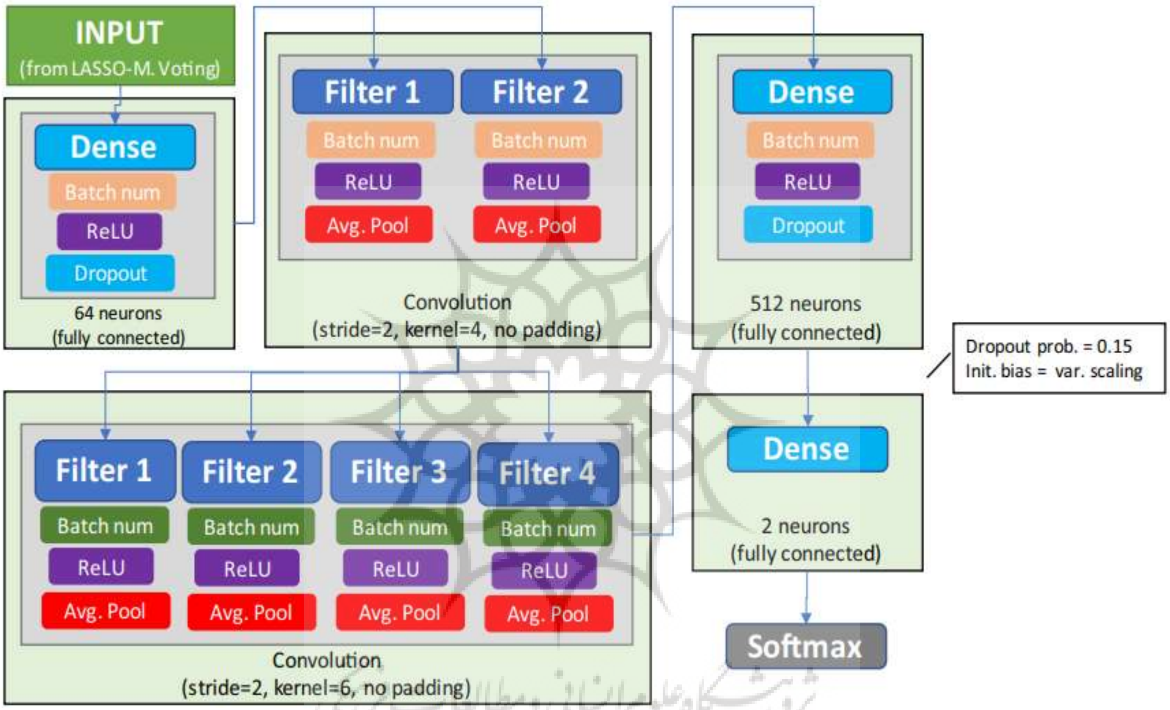

Figure 3. Proposed CNN Model

Results and Discussion

The experimental work of this paper makes use of a dataset, which we briefly detail here. As was noted previously, we employ a state-of-the-art dataset that is specifically provided. These data points are from a larger dataset produced by African medical professionals. To make our CHD prediction, we use only 14 of the available characteristics from this dataset. In Table 1, we see a catalog of algorithmic attributes, each with a brief explanation and, when relevant, a set of allowed values.

Table 1. Dataset features and description

Feature	Description
Age	Age of the person admitted
Cp	Chest pain type
Sex	NA
trestbps	Blood Pressure
Chol	Cholesterol level
Fbs	Blood Sugar
restecg	Results of electrocardiography while resting
thalach	Max heart rate
exang	Angina induced by exercise
Old peak	ST induction of depression by exercise relative to rest
Slope	The slope of the peak exercise AST segment
Ca	Number of major vessels (0–3) colored by fluoroscopy
Thal	NA
Num	Finding for the existence of heart disease, status

Performance Metrics

Some performance measures can be utilized to assess the classifier's efficacy. The effectiveness of a classifier in machine learning can be judged using a number of different metrics. The next section elaborates on a few of these standards.

Accuracy

As seen in equation 1, it is the proportion of accurately anticipated cases relative to all instances.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Where TP (True-Positive), TN (True-Negative), False-Positive (FP) and False-Negative (FN)

Precision

Equation 2 can be used to calculate precision, which is the measure of accuracy used to assess a classifier's efficacy. If the precision is great, then the number of false positives is also low. False positives increase as the model's precision decreases.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

Recall

It's a metric for assessing the classifier's comprehensiveness, and it's calculated with equation 3. When recall is strong, false negatives are low and vice versa when recall is low. In most cases, raising recall means lowering precision.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

F-Score

F-score, a measure of precision &recall, is determined by the pursuing equation4.

$$\mathrm{F1-score} = \frac{\mathrm{Recall} \times \mathrm{Precision}}{\mathrm{Recall} + \mathrm{Precision}} \tag{4}$$

Heart Disease Classification

After the data set has been presented as a feature matrix, it will be partitioned into the appropriate categories. In the first scenario, the dataset was split into two groups: those with heart disease and those without. Next, the data is classified into binary classifications using CNN, and the resulting classifier model is saved to disk. Binary classification accuracy was calculated to be 97%. The suggested model's precision, recall, F1-score, and accuracy are shown in Table2 and figure4. Table2 shows, the proposed design have an overall accuracy of 98%. Another classifier was developed, and it too was trained on four classes to represent the various heart disease conditions found in the dataset. To count all possible classifications, we ran a program developed in Python. The software examined the information, then sorted it into four categories: diseases of the first, second, third, and fourth types. The job robotically classified the records into the most closely associated class after the number of classes was calculated. The results of testing the trained model across four classes are shown in Table 3 and figure 5. With 4 classes, the expected accuracy is 87%, which is better than the data shown. The performance assessment of the design with 4 categories such as Precision, Recall, F1-score, and Accuracy values are shown in Table 3.

Table 2. Binary Classifier performance

Performance Metric	Value (%)
Accuracy	98
Precision	97
Recall	96
F1-score	98

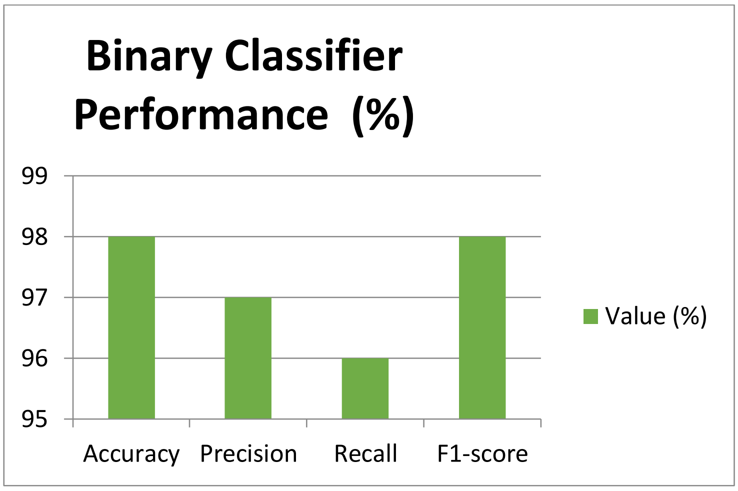

Figure 4. Binary Classifier Performance

Table 3. Multi Classifier performance

Performance Metric	Value (%)
Accuracy	87
Precision	87
Recall	82
F1-score	85

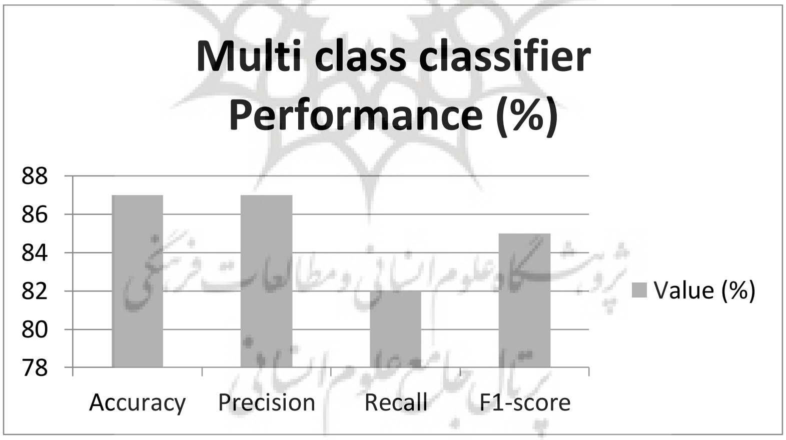

Figure 5. Multi-class classifier Performance

Comparison with Other Techniques

The suggested approach offers using the learned classifier for evaluating dataset prediction. The accuracy results of the proposed framework are compared with those of other works in Table 4, and the related figure is shown in Figure 6. The figure above shows that the proposed method has higher performance, measured in terms of total accuracy, than the other methods considered. Similarly, the decision tree-based strategy presented is similar to the one proposed here.

Table 4. Performance evaluation of the proposed model with other works

Model Name	Accuracy (%)
Risk factor based approach	86.7
Decision Tree	95.6
ANN based	94.2
Rmonto ontology	90
Proposed Model	98

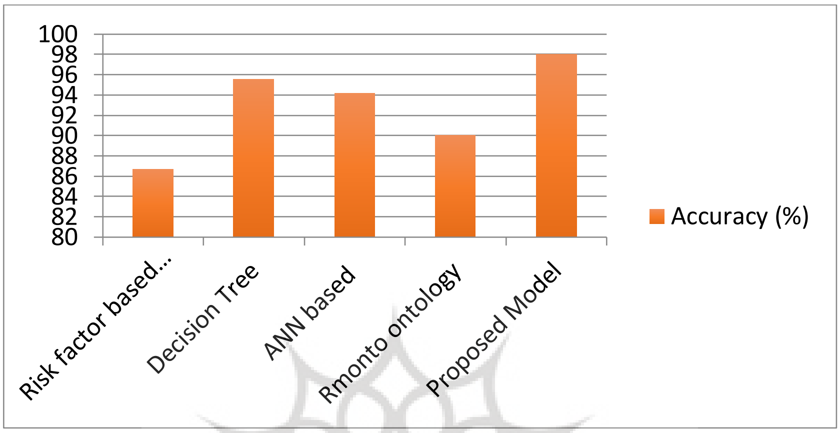

Figure 6. Proposed model accuracy comparison with existing models

The ROC curve is a graphical representation of the classifier's performance. Figure 7 depicts the true-positive-rate (TPR) vs the false-positive-rate (FPR) at a variety of classification thresholds.

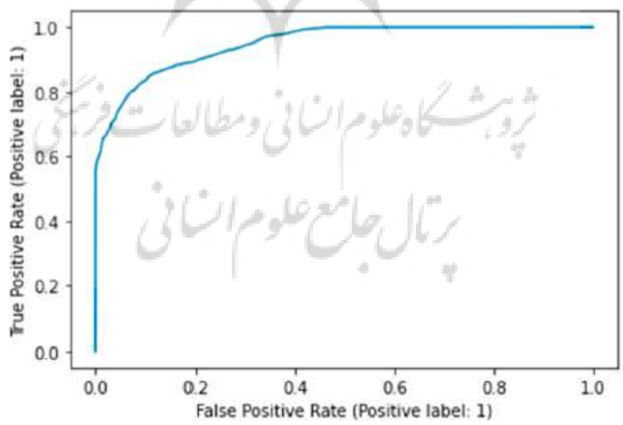

Figure 7. ROC under curve of CNN

Conclusion

Deaths from heart attacks could be avoided if CVD could be anticipated at an earlier stage. The existence of cardiovascular disease can be predicted in advance by the physician thanks to a solid classification system. In this study, we employ an available traditional dataset from the UCI repository and convolutional-neural-networks (CNNs) to make predictions on the likelihood of cardiac disease. Some parameters from heart tests and general human habits are included in this data set. The outcomes demonstrate that the proposed model is superior to the prior art methods used in this work. The proposed model has a 98% overall accuracy. In future, expand this study's applicability to the forecasting of other serious diseases, such as cancer and neurological disorders.

Conflict of interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Ahmed, S. T. (2017, June). A study on multi objective optimal clustering techniques for medical datasets. In *2017 international conference on intelligent computing and control systems (ICICCS)* (pp. 174-177). IEEE.

Ahmed, S. T., Kumar, V. V., Singh, K. K., Singh, A., Muthukumaran, V., & Gupta, D. (2022). 6G enabled federated learning for secure IoMT resource recommendation and propagation analysis. *Computers and Electrical Engineering*, *102*, 108210.

Alizadeh, F., Morell, E., Hummel, K., Wu, Y., Wypij, D., Matthew, D. & Blume, E. D. (2022). The surprise question as a trigger for primary palliative care interventions for children with advanced heart disease. *Pediatric Cardiology*, *43*(8), 1822-1831.

Almustafa, K. M. (2020). Prediction of heart disease and classifiers' sensitivity analysis. *BMC bioinformatics*, *21*(1), 1-18.

Ansarullah, S. I., Saif, S. M., Kumar, P., & Kirmani, M. M. (2022). Significance of visible non-invasive risk attributes for the initial prediction of heart disease using different machine learning techniques. *Computational intelligence and neuroscience*, *2022*.

Ashreetha, B., Devi, M. R., Kumar, U. P., Mani, M. K., Sahu, D. N., & Reddy, P. C. S. (2022). Soft optimization techniques for automatic liver cancer detection in abdominal liver images. *International journal of health sciences*, *6*.

Baskar, S., Nandhini, I., Prasad, M. L., Katale, T., Sharma, N., & Reddy, P. C. S. (2023, November). An Accurate Prediction and Diagnosis of Alzheimer's Disease using Deep Learning. In 2023 IEEE North Karnataka Subsection Flagship International Conference (NKCon) (pp. 1-7). IEEE.

Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S., & Singh, P. (2021). Prediction of heart disease using a combination of machine learning and deep learning. *Computational intelligence and neuroscience*, *2021*.

Budholiya, K., Shrivastava, S. K., & Sharma, V. (2022). An optimized XGBoost based diagnostic system for effective prediction of heart disease. *Journal of King Saud University-Computer and Information Sciences*, *34*(7), 4514-4523.

Chaurasia, D. V., & Pal, S. (2013). Early prediction of heart diseases using data mining techniques. *Caribbean journal of Science and Technology*, *1*, 208-217.

Chitra, S., & Jayalakshmi, V. (2022). Prediction of heart disease and chronic kidney disease based on internet of things using RNN algorithm. In *Proceedings of Data Analytics and Management: ICDAM 2021, Volume 1* (pp. 467-479). Springer Singapore.

Doppala, B. P., Bhattacharyya, D., Janarthanan, M., & Baik, N. (2022). A reliable machine intelligence model for accurate identification of cardiovascular diseases using ensemble techniques. *Journal of Healthcare Engineering*, *2022*.

El-Hasnony, I. M., Elzeki, O. M., Alshehri, A., & Salem, H. (2022). Multi-label active learning-based machine learning model for heart disease prediction. *Sensors*, *22*(3), 1184.

Faieq, A. K., & Mijwil, M. M. (2022). Prediction of heart diseases utilising support vector machine and artificial neural network. *Indonesian Journal of Electrical Engineering and Computer Science*, *26*(1), 374-380.

Gavhane, A., Kokkula, G., Pandya, I., & Devadkar, K. (2018, March). Prediction of heart disease using machine learning. In *2018 second international conference on electronics, communication and aerospace technology (ICECA)* (pp. 1275-1278). IEEE.

Huang, J., Li, Z., Zhang, W., Lv, Z., Dong, S., Feng, Y. & Zhao, Y. (2023). Explainable machine learning-assisted origin identification: Chemical profiling of five lotus (Nelumbo nucifera Gaertn.) parts. *Food Chemistry*, *404*, 134517.

Kumar, K., Pande, S. V., Kumar, T., Saini, P., Chaturvedi, A., Reddy, P. C. S., & Shah, K. B. (2023). Intelligent controller design and fault prediction using machine learning model. *International Transactions on Electrical Energy Systems*, *2023*.

Latha, C. B. C., & Jeeva, S. C. (2019). Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. *Informatics in Medicine Unlocked*, *16*, 100203.

LK, S. S., Ahmed, S. T., Anitha, K., & Pushpa, M. K. (2021, November). COVID-19 outbreak based coronary heart diseases (CHD) prediction using SVM and risk factor validation. In *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)* (pp. 1-5). IEEE.

Lokesh, S., Priya, A., Sakhare, D. T., Devi, R. M., Sahu, D. N., & Reddy, P. C. S. (2022). CNN based deep learning methods for precise analysis of cardiac arrhythmias. *International journal of health sciences*, *6*.

Mamatha, B., Rashmi, D., Tiwari, K. S., Sikrant, P. A., Jovith, A. A., & Reddy, P. C. S. (2023, August). Lung Cancer Prediction from CT Images and using Deep Learning Techniques. In 2023 Second International Conference on Trends in Electrical, Electronics, and Computer Engineering (TEECCON) (pp. 263-267). IEEE.

Manimurugan, S., Almutairi, S., Aborokbah, M. M., Narmatha, C., Ganesan, S., Chilamkurti, N., & Almoamari, H. (2022). Two-stage classification model for the prediction of heart disease using IoMT and artificial intelligence. *Sensors*, *22*(2), 476.

Muser, D., Liang, J. J., Castro, S. A., Lanera, C., Enriquez, A., Kuo, L. & Santangeli, P. (2019). Performance of prognostic heart failure models in patients with nonischemic cardiomyopathy undergoing ventricular tachycardia ablation. *JACC: Clinical Electrophysiology*, *5*(7), 801-813.

Muthappa, K. A., Nisha, A. S. A., Shastri, R., Avasthi, V., & Reddy, P. C. S. (2023). Design of high-speed, low-power non-volatile master slave flip flop (NVMSFF) for memory registers designs. *Applied Nanoscience*, 1-10.

Nancy, A. A., Ravindran, D., Raj Vincent, P. D., Srinivasan, K., & Gutierrez Reina, D. (2022). Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning. *Electronics*, *11*(15), 2292.

Reddy, P. C. S., Sucharitha, Y., & Narayana, G. S. (2021). Forecasting of Covid-19 Virus Spread Using Machine Learning Algorithm. *International Journal of Biology and Biomedicine*, *6*.

Riyaz, L., Butt, M.A., Zaman, M. and Ayob, O., 2022. Heart disease prediction using machine learning techniques: a quantitative review. In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2021, Volume 3* (pp. 81-94). Springer Singapore.

Sampath, S., Parameswari, R., Prasad, M. L., Kumar, D. A., Hussain, M. M., & Reddy, P. C. S. (2023, December). Ensemble Nonlinear Machine Learning Model for Chronic Kidney Diseases Prediction. In 2023 IEEE 3rd Mysore Sub Section International Conference (MysuruCon) (pp. 1-6). IEEE.

Saqlain, S. M., Sher, M., Shah, F. A., Khan, I., Ashraf, M. U., Awais, M., & Ghani, A. (2019). Fisher score and Matthews correlation coefficient-based feature subset selection for heart disease diagnosis using support vector machines. *Knowledge and Information Systems*, *58*, 139-167.

Shanmugaraja, P., Bhardwaj, M., Mehbodniya, A., VALI, S., & Reddy, P. C. S. (2023). An Efficient Clustered M-path Sinkhole Attack Detection (MSAD) Algorithm for Wireless Sensor Networks. *Adhoc & Sensor Wireless Networks*, *55*.

Song, C., Liu, B., Yang, D., Diao, H., Zhao, L., Lu, Y. & Zhang, X. (2015). Association between interleukin-6 gene− 572G> C polymorphism and coronary heart disease. *Cell biochemistry and biophysics*, *71*, 359-365.

Suneel, S., Balaram, A., Amina Begum, M., Umapathy, K., Reddy, P. C. S., & Talasila, V. (2024). Quantum mesh neural network model in precise image diagnosing. Optical and Quantum Electronics, 56(4), 559.

Bibliographic information of this paper for citing:

Kumar Gopu, Krishna Lava & Kannan, Suthendran (2024). An Accurate Prediction Framework for Cardiovascular Disease Using Convolutional Neural Networks. *Journal of Information Technology Management,* 16 (1), 27-43. https://doi.org/ 10.22059/jitm.2024.96373