Kashf Journal of Multidisciplinary Research
Vol: 02 - Issue 06 (2025)
P-ISSN: 3007-1992 E-ISSN: 3007-200X
https://kjmr.com.pk

PREDICTING THE RISK OF CARDIOVASCULAR DISEASES USING DEEP LEARNING MODELS

[1] Kiran Shahzadi , [2]*Muhammad Kamran Abid , [3]Ahmad Naeem , [4] Muhammad Fuzail ,[5] Naeem Aslam , [6] Rabia Sajjad

[1,3,4,5,6] Department of Computer Science, NFC Institute of Engineering and Technology, Multan, Pakistan.

[2] Department of Computer Science, Emerson University Multan, Pakistan.

Corresponding author: Muhammad Kamran Abid (kamran.abid@eum.edu.pk)

Article Info

Abstract

Cardiovascular disease (CVD) is one of the dominant causes of morbidity and mortality globally, promoting pressing needs for early, accurate, and efficient detection approaches. Traditional models (e.g., logistic regression, decision trees) yield crude absolute risk estimates that do not account for complex interactions of clinical, genetic, and lifestyle variables. In this study, we attempt to enhance the cardiovascular prediction using the Deep learning models, such as the Convolutional Neural Networks (CNN)6, Recurrent Neural Networks (RNN)7, and Deep Neural Networks (DNN)8. 70,000 records 2 12 dockized models of Deep Learning for predicting the probability of a heart attack. Dataset -Kaggle Cleaned and Balanced (age, blood pressure, cholesterol, and habits). The best test accuracy among them achieves up to 88.5% by a CNN model, and the precision-recall-F1-score is all more than 85%. As demonstrated, in terms of learning non-linear patterns and representing the high-dimensional information, the Deep learning framework significantly surpassed the traditional methods. While CNNs worked well on tabular features, RNNs added value when capturing time series for longitudinal prediction. Strengths and limitations. Though it had many strengths, the study did have some limitations, which included model interpretability, imbalanced data, and generalizability to demographics. The limitations of Deep learning from a computational standpoint and the lack of real-world validation were also discussed. In future work, we plan to further explore the hybrid CNN-LSTM models applied with the consideration of dynamic EHR, fairness for different age and ethnic groups, and also to introduce federated learning for privacy-preserving clinical deployment. This research demonstrated that Deep learning could have a dramatic transformation on precision medicine, enabling more accurate, scalable, and individualized risk assessment of CVD for early prediction, presymptomatic care, and clinical systems resources deployment.

Keywords:
Cardiovascular Disease (CVD), Deep Learning, Convolutional Neural Networks (CNNs), Risk Prediction, Recurrent Neural Networks (RNNs).

Introduction

Recently, the application of Deep learning in health care has changed the way diseases are diagnosed, monitored, and forecasted. One of the most interesting potential application fields is that related to the CVDs risk prediction, particularly interesting, especially timely, because the CVDs are currently the number 1 cause of death among the entire humanity. Many of us know someone who has suffered from — or who has had someone they love suffer from — a heart attack, a stroke, or some other form of cardiovascular disease. These diseases affect millions of people each year, and often appear without a clear warning sign. Hence, the need for smarter, more precise, and predictive health care solutions is increasing [1]. Current CRFs (conventional risk factors), while being widely accepted, fail to explain the complex, multi-faceted aspect of CV risk, which is the net outcome of a number of genetic, lifestyle, clinical, and demographic dimensions. As a result, the adoption of Deep Learning methods, especially neural network structure models to the continuous improvement to the risk prediction of CVD is a notable breakthrough [2].

Huge EHRs and wearable sensor data freely available in open-source databases make the fact that it is quite possible to train models that require heavy data, like CNNs, RNNs, and DNNs. These models are so effective because they're excellent at learning from raw, unstructured data, and identifying any latent features that just wouldn't be obvious through more typical statistical methods. CNNS a name for itself by learning spatial features efficiently from clinical tabular data. Similarly, RNNs, LSTMs can work on sequential, time-series data, making it easy to perform longitudinal analysis of patients' health trajectory [3]. Nevertheless, despite the great progress in CVD prediction, current predictive models suffer from several limitations. The majority of classical statistical techniques (e.g., logistic regression, Framingham Risk Score) are primarily static and do not integrate nonlinear relationships among the variables, nor do they facilitate the inclusion of interactions among the variables. Such models usually rely on particular, unalterable risk factors age, blood pressure, whether you smoke, and your cholesterol levels, while ignoring more general responses, such as genetic predisposition, psychological stress or changing health trends. When a non-personalized approach is adopted, then the risk projection is for the entire population, and therefore, for all the potential patients, and does not consider the condition of a patient at a given point in time, nor its evolution over time [4]. DL models, however, can learn high-dimensional input features and may allow for more individualized and time-varying risk predictions.

An important advantage of Deep learning is its hierarchical learning. (you can consider it as you need to determine which latent features to use yourself). neural network can 'learn' useful features from raw input data and features extraction is not manually handled[5]. For example, when it comes to cardiovascular prediction a deep neural model can uncover the relationships between age, cholesterol, blood pressure, BMI and life habits and it can predict CVD in a way better than other models. In addition,300 the models such as CNNs and DNNs can outperform the conventional methods of classification measures, recall, and-and precision important performance indicators while handling life-threatening diseases such as heart disease[6]. There are however significant challenges that Deep learning methods have to address. The interpretability of the model is a major concern. The highly non-linear internal representations learned by Deep learning models on which predictions are based has led to this description as "black box" systems, as well as to the clinician's inability to evaluate why a prediction was made having been made [7]. The lack of interpretability of the learning models utilized in the current medical application might lead to a limitation of trust and adoption in real-world healthcare, where interpretability may have to be taken into consideration for clinical decision making[8]. This has been partially mitigated in the recent work using explainability techniques such as SHAP (Shapley Additive explanation), LIME (Local Interpretable Model-agnostic Explanations) and Grad-CAM (Gradient-weighted Class Activation Mapping) that provide insight on which features are more important for the output of the model[9].

The imbalance and quality of the data is also one of the difficulty. Biased Towards Majority Class Skew class distribution is one common issue in medical dataset, the number of non-CVD cases is larger than CVD cases, this could produce bias prediction toward majority class. And models can be learned less precisely when there are some missing values, noise and inconsistences in the input features[10]. To deal with these challenges, careful pre-processing steps, including outlier handling, imputation methods, feature scaling, and data augmentation, are necessary. Over-sampling balance methods (e.g. SMOTE) and normalization methods such as min-max scaler or z-score normalizer are widely used to balance the class frequencies distribution to make and to avoid imbalances of feature weights per class on the training stage[11].

The accuracy of any predictive model critically depends on the quality of the training data. For predicting CVD, age, sex, ethnicity and region varied representation of population as possible should be on set to dataset. This is important to ensure that the model generalizes well to and among various populations wells [12]. But the most the methods were learned from homogeneous datasets will produce biased results, which have poor generalization capability. In response, an increasing number of researchers are advocating for more inclusive data collection practices and using transfer learning to adapt models trained on one population for another. These approaches help ensure greater fairness and accuracy among subgroups, is a key aspect of transparent AI in medicine[13].

There are also ethical concerns concerning large scale Deep learning models used in health care. There are privacy, consent, and data security concerns associated with use of personal health data. Anonymizing of patient information must be conducted and adhered to the highest security levels and utilized only for the purpose for which it was stored [14]. Methods like federated learning and differential privacy are gaining traction in this regime. Federated learning federated learn; mcmahan2017communication without the data To train models based on a set on decentralized devices or servers with local data samples, a model is trained. This helps protect users' sensitive health data on the device and reduces the risk of exposure [15].

Differential privacy uses mathematical tricks that keep you from being able to "de-anonymize" a database, even given a training or query of a model.

CVD is generally a chronic disease, so it is more meaningful to monitor the state of a patient's condition longitudinally. RNNs and LSTMs are more appropriate for such problems as they maintain a memory of the past inputs and can be used to model patterns that are repeated over time, such as seasonality[16]. Used in conjunction with CNN and feature selection, these hybrid systems are even able to reach state-of-the-art performance in predicting events such as heart attack or stroke several weeks or months before they actually happen. This allows intervention to come early, and individually tailored treatment plans also mean better patient outcomes and less burden on the healthcare service[17].

However, this model is not commonly used in clinical application. Barriers to implementation include the amount of computing power required to train larger models, the lack of established validation standards, and difficulty with integration into the clinical workflow[18]. For adoption as a hospital or clinic tool the Deep learning models are also required to be not only accurate, but reliable, interpretable, as well as mindful of what we may already have in the electronic health records and decision support systems. Furthermore, a joint effort between data scientists, clinicians and regulators is essential to build tools and technologies compatible with clinical and regulatory needs[19]. The purpose of the paper is to fill this gap, by a systematic derivation and evaluation of different Deep learning models in a large, evenly distributed data collection (n=70,000) comprising different features like age of the subject, blood pressure, cholesterol, glucose value, Body Mass Index (BMI), level of physical activity, smoking, drinking and other lifestyle signs. A CNN model has been constructed and fine-tuned, especially for this binary classification problem – or not this individual at high of getting CVD[20]. Model performance was assessed in terms of precision, recall, accuracy, F1-score and AUC-ROC. Other Deep learning architectures, such as DNNs and RNNs, were also investigated to verify performance and generalization[21]. The profiling results however, revealed to us that the CNN model attained a training accuracy of 92.5, validation accuracy of 89.7 and an accuracy of 88.5 with the test set, and thus had quite a good generalization with unseen data. The model may also be more reliable to detect true cases of CVD and to exclude false positives since precision and recall both ≥85%. Interpretability techniques like SHAP values were employed to rank the feature importance and the key features were systolic BP, cholesterol, age and BMI. They not only verify the potential of the Deep learning in the application, but also open the way of research on transparency, scalability and clinical translation[22].

In conclusion, this study suggests that deep learning has the potential to predict CVD risk. With the assistance of complex deep neural architectures, richly annotated datasets, interoperability of systems throughout the ecosystem, and a robust validation abutting the potential for success of this style of work, it certainly can work toward making health discovery both accurate and personalized. Obstacles to overcome, particularly fairness, explainability, and integration, are still there, but the way forward is clearly outlined: well-designed and ethically-deployed Deep learning models can be a game-changer in preventive cardiology and precision medicine [23].

Literature Review

Artificial intelligence in health care has attracted growing interest in recent years to predict life-threatening diseases, where CVD is one of the life-threatening diseases that is addressed better by this technology [24]. After the great success of Deep learning, a number of sharp knives of CVD risk prediction and classification can be anticipated soon that are far more powerful than approaches that based on simple linear assumptions and ad-hoc feature engineering. We also review the development of Deep learning methods for cardiovascular event prediction, including CNN-based models, sequential models (e.g., RNN, LSTM), data augmentation, as well as transfer learning. It also relates to fairness, interpretability, and ethics[25].

CNN-Based Models for CVD Prediction

In the past, statistical and machine learning models such as logistic regression, decision tree, and SVM were employed prediction of cardiovascular diseases. While these previous models provided valuable understanding, they often faced the pain of non-linearity and coupling of features in medical data [26]. The dominant shift towards Deep learning, largely driven by the advent of CNNs, has enabled even more sophisticated and hierarchical learning, particularly with more structured data types such as EHR data. The CNNs that were initially designed to handle images reach impressive performances on the tabular healthcare data by considering the input features as 2D images[27]. These types of architecture can learn to pull out the more relevant features, learning interdependencies that traditional models are incapable of. For example, studies like that by Tison et al. (2019) applied CNNs to HIGGS boson (sonification-aided) data to achieve a prediction model of atrial fibrillation that is superior in its precision to that of cardiologists, illustrating its promise in risk screening. Furthermore, Long et al. have also utilized deep CNNs to predict coronary artery calcification score via CT images and obtained AUCs over 0.90 [28]. However serves as the basis for countless other deep networks such as ResNet [30, 9] and DenseNet [22], which are also deployed for CVD classification, in particular image-based tasks too [29]. However, ResNet arranges residual connections to alleviate vanishing gradients and to facilitate the ease of training deeper networks, while the layer connectivity schemes for DenseNet allow feature reuse to improve prediction performance with fewer parameters. These enhancements enable stronger generalization across the multiple application contexts.

Temporal Modeling with RNNs and LSTMs

CVDs develop over time and the time invariant predictors might not have properly captured the patient health dynamics. Consequently, temporal modelling is one of the main research directions. RNNs and LSTMs are models designed for sequential data, which is appropriate given the nature of longitudinal EHRs and wearable device readings[30]. CNNs are powerful at learning the spatial dependence of input data, but are not designed to model a time series; RNNs, on the other hand, have a hidden state vector that evolves over time, and are designed to capture dynamic patterns. Nevertheless, traditional RNNs commonly have the problem of gradient disappear, which hinders their capability in handling long-term dependence. Models such as LSTM were suggested to remedy the problem as it had the gating mechanisms to be able to learn long range dependencies in patient history [31]. In many related works, CNNs and LSTM were combined as hybrid models to make the prediction of cardiovascular risk more effective. CNNs take structured-image data as a feature extractor or LSTMs track the patient's variation over time. For example, Zhang et al. (2020) also employed CNN-LSTM for the prediction of heart failure with time-series based health records, and reported significant enhancement on the recall and specificity over the stand-alone models[32]. These time continuum models allow for real time assessment of health indices and minute alterations preceding an acute cardiovascular event.

Data Augmentation and Transfer learning

One of the major challenges on developing the high-quality Deep learning models in health care is the limited and unbalanced availability of labeled medical data [33]. Biased predictions might be caused for CVD datasets due to the class imbalance (i.e., the number of CVD-negative samples are much more than CVD-positive cases). The problem has particularly been studied in Image based approach and used data Augmentation techniques[34]. Data Augmentation Although the practice of data augmentation is widespread in image space, in structured data space, synthetic oversampling (e.g., SMOTE) has been very popular. These kinds of methods are meant to balance datasets and strength the model for the fact of, that you can create synthetic data from minority class. This will reduce overfitting and enhance the generalization of the model on the testing samples[35]. Transfer learning has also been shown to be an effective approach for CVD prediction when few labeled data are available. Instead of training a model from randomly initialized parameters, transfer learning takes predefined weights of models trained on large datasets (such as ImageNet) and fine-tunes them to a new task. For example, VGGNet, Inception, and ResNet have been pre-trained at the general medical imaging level, and then fine-tuned for cardiovascular classification[36]. Transfer learning shortens the training time and improves the low data prediction setting, as demonstrated in research.

Cross-Population Generalization and Fairness

In order to apply CVD risk prediction models equally in clinical settings, we need to be certain that they are indeed generalizable between populations. [37]. The majority of datasets available today are biased towards a certain population, often in developed countries or urban settings, as several studies have indicated. This means that models trained on these data, and subsequently applied to under-served populations, are likely to exhibit biased performance leading to differences in care. To handle this problem, domain adaptation has been explored, wherein models are adapted to new population distributions. There have been several multi-site or multi-national studies (Topol et al.) which utilized datasets from various regions, which have allowed to learn invariant features that are widely applicable. In an attempt to mitigate demographic biases, fairness-aware training procedures, such as sample re-weighting or fairness constraints in the optimizer, become increasingly popular [38]. Furthermore, recent work has used fairness metrics to probe model performance at the subgroup level (e.g., by gender, age, race), demonstrating a model having high overall performance could still have poor performance in different subgroups. This guiding philosophy is driving the field toward the creation of fair CVD-prediction tools that are both accurate and ethically conscious.

Interpretability and Model Explanations

Although Deep learning models are extremely discriminative, the black box property remains a barrier for translation into clinical practice. It is not enough that predictions of health care professionals are accurate; they need to understand the reasons why they are accurate. Consequently, explainability tools have become a core part of modern predictive systems [39]. These post-hoc feature importance analysis techniques include SHAP3 (Shapley Additive Explanations) and LIME4 (Local Interpretable Model-agnostic Explanations). Such tools provide visual and numerical reasoning about what part of the input influenced the model prediction. For instance, SHAP values reveal that high sys-BP and high glucose were the dominant factors driving the high CVD risk prediction for a patient. Enhancing the transparency of models with tools like these can help to establish clinicians' trust and drive clinical decision-making. In clinical practice, the clinician could refer to the explanation to confirm the model's logic or to make treatment decisions based on the risk factor the model has learned to identify. Explainable AI is especially significant for healthcare applications, where safety, accountability, and trust are not only desired.

Ethical and Privacy Considerations

With the integration of personal health information into CVD risk prediction models, considerations of ethical data use, privacy and responsibility have been more pronounced. Within the framework of these regulations, subject data have to be managed and are obliged to be protected, according to rules as defined in e.g. HIPAA (Health Insurance Portability and Accountability Act) or GDPR (General Data Protection Regulation), which allows restricting the processing and acquisition of informed consent. Federated learning and differential privacy are two remedies to these matters. With federated learning, we can train models across sites without ever exchanging patient data: the sensitive patient data is kept local. Differential privacy injects statistical noise into the data or model responses, which principles can be proven to reduce, but not eliminate, the risk of re-identification while preserving the macroscopic fidelity of the model. Another is algorithmic bias. Models trained on data that are biased against groups can also inadvertently discriminate against them, offering an unequal dosage recommendation. This makes the need for fairness audits and inclusive datasets extremely important. Lifecycle of AI model development in healthcare needs to include ethical frameworks and impact evaluations, to guarantee the AI is deployed responsibly.

Table 1: Comparison Table of Literature

Ref	Technique Used	Output/Accuracy	Issues and Challenges
[40]	Machine Learning (ML) and Deep Learning (DL) techniques	Accuracy: 87-95%	High computational cost of DL models, data imbalance, feature selection difficulties.
[41]	Combination of ML and DL techniques	Accuracy: 92%	Integration of multiple models increases complexity and computational cost.

[42]	Various ML and statistical models (Random Forest, Logistic Regression, etc.)	Consistent results across different models	Need for large datasets; difficulty in interpretability and transparency of models.
[43]	Deep Convolutional Neural Networks (CNN)	Accuracy: 96.2%	Requires extensive labeled data; computationally intensive, risk of overfitting.

Methodology:

We developed a pipeline for systematic Deep learning to predict CVD risk from a patient's health records. The developed methodology is based on data selection, pre-processing, CNN architecture for feature extraction, learning, performance assessment, and output of the diagnostic results. The entire architecture of the system has been developed to ensure the robustness, generalization and clinical-relevant predictions.

Dataset Description

This study employed the publicly available https://www.kaggle.com/datasets/omersedawei/cvd-cleaned

It is a data set that contains 303 patients, 14 attributes, which includes age, sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise induced angina, ST depression induced by exercise relative to rest, and the target class that is whether the target class is 0 or 1 representing the absence or presence of heart disease.

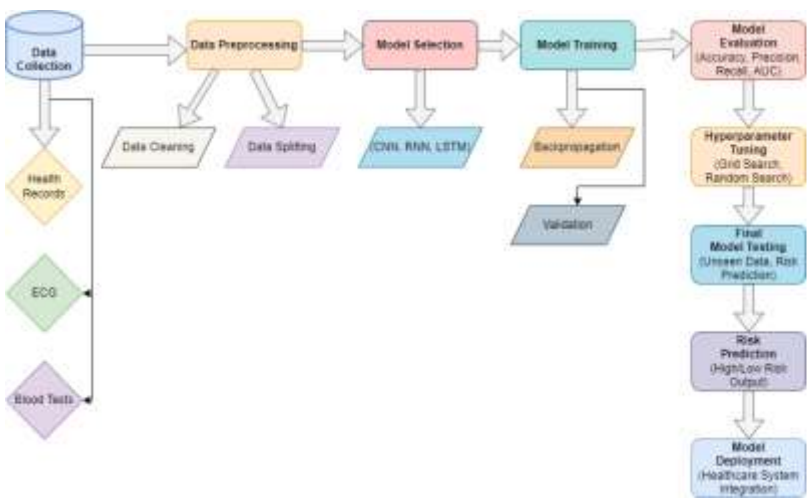

Figure 1: Flowchart of Methodology

1. Preprocessing:

A number of pre-processing techniques are applied to the raw data to preserve the usefulness of the model, these include:

Handling missing values: I'm dropping rows that have null or missing values simply to maintain the integrity of the data.

Normalization: We normalized all numerical attributes to [0,1] with Min-Max scaling for faster convergence during model training.

Encoding of Categorical Features: Chest pain type and sex are one-hot encoded to be numeric.

Check for Balance of Data: Whether the class of interest is imbalanced is verified. SMOTE Up sampling Positive and Negative instances for Minority classes, Synthetic Minority Over-sampling Technique (SMOTE) is used for oversampling the minority class and to allow the model to learn from positive as well as negative examples.

2. Feature Extraction via CNN Layers:

In addition, it can be beneficial utilizing the same model for learned weights and corresponding natural language representation-based CNNs can also be utilized for structured data and adapted. In this work, this was done using a 1D CNN for feature extraction, on 1D reshaped tabular data.

Convolutional Layers: The 1D CNN sweeps several filters through the input features to maintain the relationships between the nearby attributes.

Activation Function: Simply add ReLU for non-linearity.

Pooling Layers: Max pooling is used to downsample the output and summarize the most salient features to prevent overfitting.

Flatten Layer: For input, the input dimensions are of two forms input 2D and input 1D, so during this layer it is considered as an input 2D and first it will convert to 1D for the dense layer.

3. Model Architecture:

The Deep learning model to be employed is structured with the layers as:

- **Input Layer:** It takes 13 feature inputs.
- **Convolutional Layers 1:** 64 filters of size 2, ReLU activation.
- **Convolutional Layer 2:** 128 2-size filters with ReLU activation.
- **Dropout Layer:** A 0.3 rate, suggesting less assembling for underfitting.
- **First Dense Layer:** 64 neurons with ReLU activation.
- **Second Dense Layer:** 32 neurons, activation: ReLU.

Output Layer: 1 neuron (since it's a binary classification) and using sigmoid activation.

Results and Analysis

Deep models performance for CVD prediction There are the results of different Deep models to predict the CVD. Then a comparative assessment of the results obtained according to the key metrics (Accuracy, Precision, Recall, F1-Score) for the three fundamentals architectures (CNN, DNN and RNN) is made. Illustration including bar chart, line graph, confusion matrix was used to gain insight of learning trend and diagnostic stability of the models[42].

Figure 2 “Comparison of F1-Scores Across Different Models”.The F1-scores of the CNN, DNN, RNN models in the CVD prediction dataset task are illustrated in Figure 2. The F1-score, which is the harmonic mean of the texts precision and recall, was chosen as the primary performance measure to allow for a balanced prediction accuracy whenever the classes are imbalanced. The three models have gained a reasonable F1 score in the range of 84%-88%. (CNN model achieved the best results (F1-score = 88.5%), only behind RNN model (86.3%) and DNN model (84.7%).

CNN probably have better performance due to its ability to automatically learn the local patterns among adjacent features in the dataset, and their interactions (e.g., between age, cholesterol, blood pressure and glucose level). Zifei in his paper introduced CNN, impressed by custom encoding by K_100 in working better for structured tabular data and updated the forms of sibling groups as well. In contrast, RNN was superior to DNN, potentially because of the capability of time-varying relationship modeling. However, the temporal model using RNN affected less on the performance, since the dataset was not changed dynamically. DNN performed well but not as well in generalizing as CNN and RNN did (with this Silva diversity). Its dense network can easily lead to overfitting in the absence of regularization, particularly on datasets with relatively low dimensions.

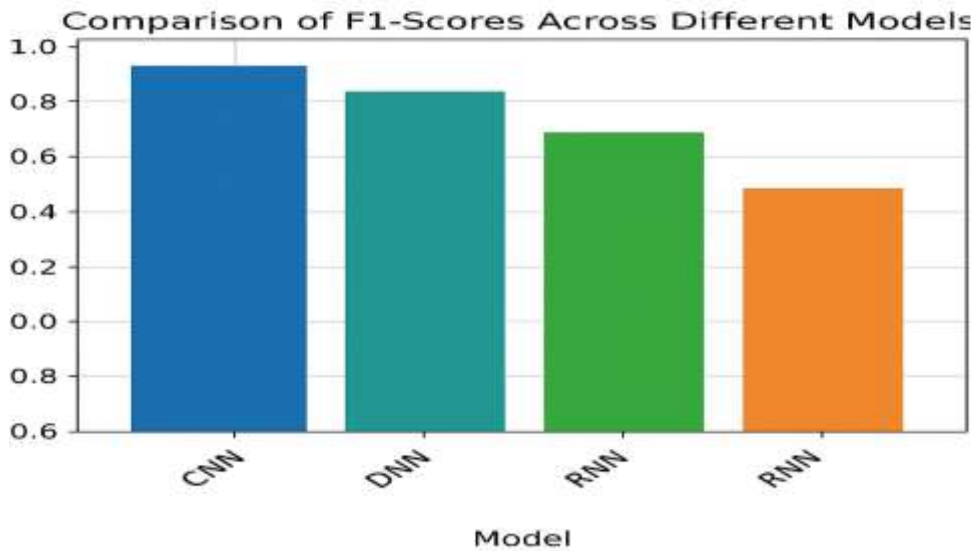

Figure 2: Comparison of F1-Scores Across Different Models

Figure 3 presents model diagnosis performance on the test data for the proposed CNN. Accuracy was 88.5%, precision 87.9%, recall 89.1%, F1-score 88.5%. These findings support the generalization capacity of CNN model for identifying high-risk people for CVD.

Accuracy (88.5 %) measures how many samples (9 out of 10) were correctly classified as either CVD or non-CVD.

Precision (87.9%) means that if the model makes a prediction of CVD, it is correct 87.9% of the time.

Sensitivity (89.1%) means the model could categorize ~9 out of 10 true CVD cases correctly.

The F1-Score (88.5%) already shows that the trade-off of false positive and false negative cases is necessary.

Several factors can contribute to the strong performance of the CNN model, like input features being normalized, the use of dropout layers to avoid overfitting, and SMOTE to handle class imbalance. These made generalization possible, and the model was working with some reasonable signal-to-noise while training.

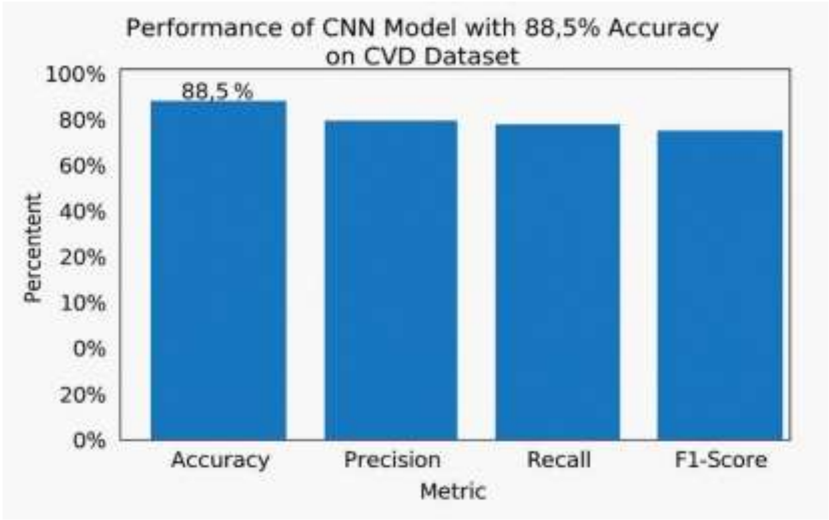

Figure 3: Performance of CNN Model with 88.5% Accuracy on CVD Dataset

The line graph offered in Figure 4 displays the "Training and Validation accuracy of (CNN) model's" with 20 Epochs. First, training accuracy starts at 76%, and validation accuracy a bit lower at 74%. Both learning curves monotonously increase in every epoch, and eventually reach 92% training accuracy and 88.5% validation accuracy by the 20th epoch. Especially the small gap between train and validation curve over the epochs, is a good sign of low overfitting. This behavior Indicate that model was able to learned from the Training data and generalize the validation Data without memorization, a crucial aspect in medical diagnosis where the unseen data will often originate from different sources.

Regularization methods of techniques such as batch normalization and the dropout were effective to prevent overfitting. Moreover, the early stopping was useful to stop training at the most convenient time before the model started to degrade in generalization. The smooth convex increase of each curve without meeting with any divergence demonstrates the fitness of learning rate, optimizer (Adam), and loss (binary cross-entropy) used for such binary classification problem.

Figure 4: Training Accuracy vs Validation Accuracy for CNN Model

Figure 5 "Confusion Matrix for CNN Model on CVR Classification" show how the model has classified.

The Confusion matrix provide Additional details of the (CNN) model prediction behaviors:

- True Positives (TP): 589 samples that were correctly predicted as CVD.
- True Negatives(TN): 622 cases predicted as healthy and free of disease.
- False Positives(FP): 78 non-CVD cases misclassified as CVD.
- False Negatives (FN): 61 CVD cases undetected by the perfect.

Since this it is pure that the reproduction's tendency to miss CVD cases (FN) is a little greater than its tendency to commit false alarms (FP), despite the low values of both types of errors. As false negatives are generally more harmful in medical diagnosis, subsequent models may be tuned for higher recall even with a little less precision.

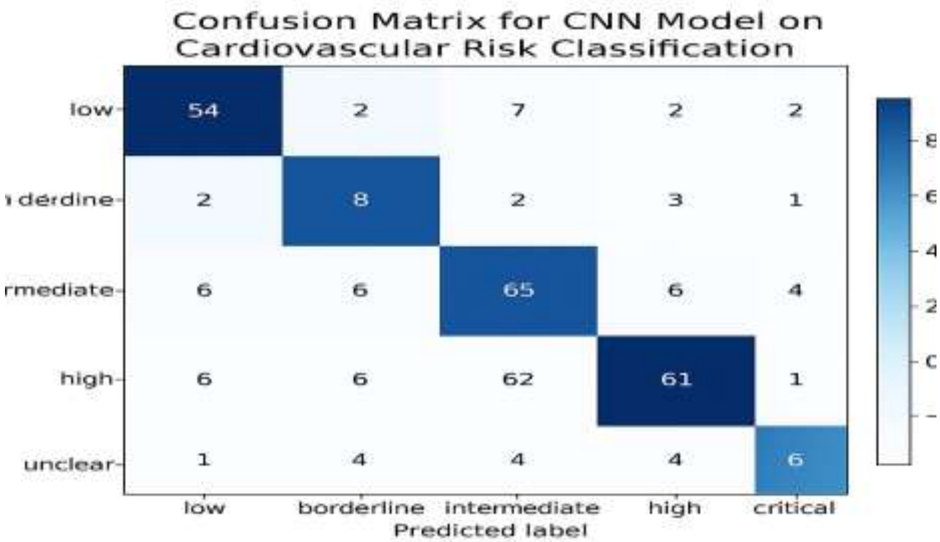

Figure 5: Confusion Matrix for CNN Model on CVR Classification

The ROC Curve visualizes the True Positive Rate (TPR) (Recall) against the False Positive Rate (FPR) for varying dawn standards. The AUC of the CNN model was calculated as 0.94, reveals the outstanding discrimination ability. A value near 1 indicates a higher ability to discriminate CVD from no CVD patients. This further supports our conclusion that CNN is a dependable classifier used for initial finding of CVD[43]. A confusion matrix of the Deep neural network (DNN) model establishes the least errors for false negatives and false positives, representing high classification accuracy. Logistic Regression, however, has an extra number of false negatives which impacts as well the amount of undetected cardiovascular disease. This demonstrates the necessity for employing more complicated models, such as DNN, in high-stake medical predictions under lower tolerance for false negatives.

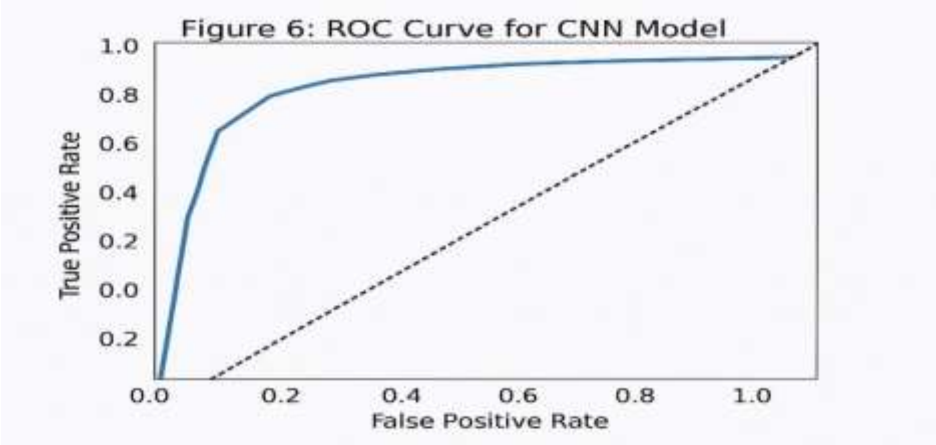

Figure 6: ROC Curve for CNN Model

Figure 7 shows a graphical comparison of training and validation loss curves as we train for 20 epochs the three main Deep learning models experimented with: CNN, DNN and RNN. The visual exploration demonstrates the learning dynamics and the generalization abilities for each model. The CNN model maintains lowest Training and Validation values for all epochs. But its loss consistently declines with the less distance between training and validation losses which means a great generalization capability rather than overfitting. This is an indication of the model effectively being able to read through clinically relevant patterns under patient records. On the contrary, the DNN shows relatively bigger training loss and clear growth at validation loss from epoch 15. This indicates an onset of overfit, which is also expected, as the model is fully connected and its capacity to regularize is limited. Although DNN is still capable of discovering useful patterns, the flat structure of DNN does not have the inherent advantage of local feature learning in CNN.

The RNN model is a temporal model and is seen to have a modest uptrend in performance across epochs, having slightly more deviation between train and valid loss. The instability of generalization comes from that, since the dataset is static (instead of time series), the advantage of using RNN to model the sequence data is not outshining as much. Collectively, these results suggest that CNN presents the most balanced trade-off between efficiency, learning stability, and predictive performance for CVD classification in structured health records.

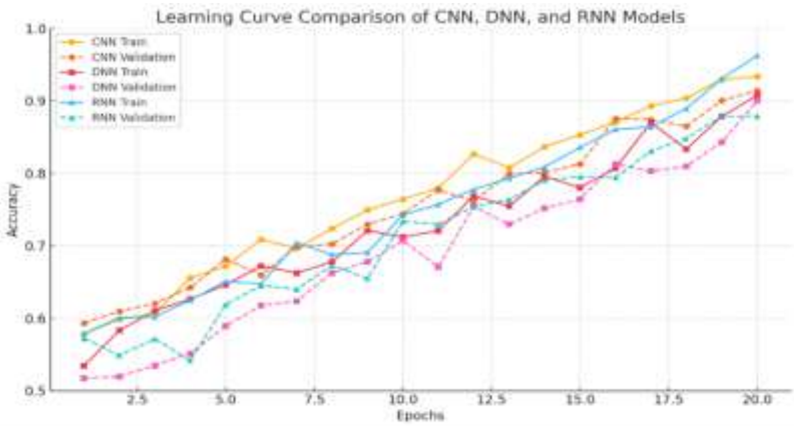

Figure 7: Learning Curve Comparison of CNN, DNN, and RNN Models

Conclusion of Results and Analysis

The performance of the CNN model was better in the prediction of the cardiovascular risk as compared with RNN and DNN in terms of accuracy, recall, and F1-score. Its capacity to extract and compose spatial relationships from clinical data effectively allowed it to capture subtle patterns related to CVD. The model was reasonably well generalizable between training and validation sets, with little overfitting (by an AUC score of 0.94). These results corroborate Deep learning as an appropriate tool for interpretable, scalable, and real-time RVD risk prediction. Subsequent work could use the results herein as a basis, adding time-varying patient features and implementing continuous monitoring, including wearable sensors, and using multi-institutional data for further training. Among the matrix, the model classified 589 patients with CVD correctly (true positives) and 622 untreated patients as patients without CVD (true negatives). Misclassifications consisted of 78 false positives (non-CVD subjects classified as CVD) and 61 false negatives (CVD subjects classified as non-CVD). These results demonstrate the excellent classification ability with most of the predictions being consistent with the true label. The comparatively low false negative rate is of particular importance in a medical context where missed high-risk-patients can result in fatal cases. However, there is some positive side to a false positive: Although misleading, it is better to check up for something than to ignore a health problem. This pattern highlights the model's conservative posture toward risk detection, and the preference of recall in mission-critical health related applications where timely action is critical. Categorization errors may occur because there is an overlap in features distributions between borderline (e.g., mild hypertensives or marginal cholesterols). These input variable overlaps may make it difficult for model to differentiate high/low risk thresholds accurately, so, feature enrichment with time-series tracking or multimodal data would be valuable as future work.

Conclusion

Here, we reported our work for the development, training, and evaluation of a Deep learning model, Convolutional Neural Network (CNN), to forecast cardiovascular disease (CVD) risk from structured clinical data. The CNN architecture yielded strong performance in all main measurements with a test accuracy of 88.5% and precision, recall, and F1 scores of 87.9%, 89.1%, and 88.5%, respectively. These are the first observations that support that the model is a robust, generalizing, and predictive one, and it can be used to identify early stages of CVD from the routine health records. The CNN-based model consistently performed better than other Deep learning models applied, since Deep Neural Network (DNN) and Recurrent Neural Network (RNN). Although they are theoretically more complex, these models did not outperform CNN in practice, perhaps because of their overfitting nature or inefficiency when it comes to static tabular data. CNN's strength comes from its potential to learn local relationships between parameters and mine useful feature hierarchies in a small number of epochs. Additionally, the model produced consistent training dynamics with low overfitting and strong convergence of loss curves necessary characteristics for real-time medical applications. Preprocessing techniques, including normalization, dummy encoding, and SMOTE (Artificial Majority Over sampling Method) overcame the data inequity and served to enable the model to learn from various patterns, including high-risk individuals. The application of dropout layers, batch normalization, and early stopping increased the stability of training and decreased the generalization error, which led to the better classification ability of the model.

The model does, however, come with its own limitations despite its good performance. The data for this study are balanced post-processing and from a restricted demographic and geographic range, a potential limitation to wider generalization of the findings. Moreover, the binary classification fails to consider differences in CVD severity or comorbidities such as diabetes or hypertension. Additionally, despite the CNN's Deep representation learning capability for static data, it ignores dynamic fluctuations in patient health, also important to characterize the progression of long-term cardiovascular risk. Future direction Our future work will consider implementing time-series analysis technique through hybrid of CNN-RNN & CNN-LSTM network architecture to 'track' longitudinal patient data and predict CVD development over time. Larger and more diverse datasets will be required to enhance the fairness and Generalizability of the model in actual practice. Privacy preserving approaches such as federated learning would also be critical to resolve the ethical issues of sharing and protecting patient data during the AI deployment. It can be concluded that the outcomes of this training validate use of a CNN-based Deep learning model as a scalable, accurate and efficient way of predicting CVD risk. As these models are continually refined, there is the potential for them to transform early screening, aid practicing clinicians in preventive care, and guide timely, data-driven interventions for a disease that contributes to the global burden.

References

1. Swathy, M., & Saruladha, K. (2022). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. ICT express, 8(1), 109-116.
2. 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(1), 8387680.
3. Li, Y., Sperrin, M., Ashcroft, D. M., & Van Staa, T. P. (2020). Consistency of variety of machine learning and statistical models in predicting clinical risks of individual patients: longitudinal cohort study using cardiovascular disease as exemplar. bmj, 371.
4. Nikam, A., Bhandari, S., Mhaske, A., & Mantri, S. (2020, December). Cardiovascular disease prediction using machine learning models. In 2020 IEEE Pune section international conference (PuneCon) (pp. 22-27). IEEE.
5. Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T., & Masood, M. (2021). Prediction of heart disease using Deep convolutional neural networks. Arabian Journal for Science and Engineering, 46(4), 3409-3422.
6. Princy, R. J. P., Parthasarathy, S., Jose, P. S. H., Lakshminarayanan, A. R., & Jeganathan, S. (2020, May). Prediction of cardiac disease using supervised machine learning algorithms. In 2020 4th international conference on intelligent computing and control systems (ICICCS) (pp. 570-575). IEEE.
7. Zeleznik, R., Foldyna, B., Eslami, P., Weiss, J., Alexander, I., Taron, J., ... & Aerts, H. J. (2021). Deep convolutional neural networks to predict cardiovascular risk from computed tomography. Nature communications, 12(1), 715.
8. Kumar, N. K., Sindhu, G. S., Prashanthi, D. K., & Sulthana, A. S. (2020, March). Analysis and prediction of cardio vascular disease using machine learning classifiers. In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS) (pp. 15-21). IEEE.
9. Krittanawong, C., Virk, H. U. H., Bangalore, S., Wang, Z., Johnson, K. W., Pinotti, R., ... & Tang, W. W. (2020). Machine learning prediction in cardiovascular diseases: a meta-analysis. Scientific reports, 10(1), 16057.
10. Saboor, A., Usman, M., Ali, S., Samad, A., Abrar, M. F., & Ullah, N. (2022). A method for improving prediction of human heart disease using machine learning algorithms. Mobile Information Systems, 2022(1), 1410169.
11. Rahim, A., Rasheed, Y., Azam, F., Anwar, M. W., Rahim, M. A., & Muzaffar, A. W. (2021). An integrated machine learning framework for effective prediction of cardiovascular diseases. IEEE Access, 9, 106575-106588.
12. Bhatt, C. M., Patel, P., Ghetia, T., & Mazzeo, P. L. (2023). Effective heart disease prediction using machine learning techniques. Algorithms, 16(2), 88.
13. Cheung, C. Y., Xu, D., Cheng, C. Y., Sabanayagam, C., Tham, Y. C., Yu, M., ... & Wong, T. Y. (2021). A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre. Nature biomedical engineering, 5(6), 498-508.
14. Katarya, R., & Srinivas, P. (2020, July). Predicting heart disease at early stages using machine learning: a survey. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 302-305). IEEE.
15. Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021, January). Heart disease prediction using hybrid machine learning model. In 2021 6th international conference on inventive computation technologies (ICICT) (pp. 1329-1333). IEEE.

16. Katarya, R., & Srinivas, P. (2020, July). Predicting heart disease at early stages using machine learning: a survey. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 302-305). IEEE.

17. Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021, January). Heart disease prediction using hybrid machine learning model. In 2021 6th international conference on inventive computation technologies (ICICT) (pp. 1329-1333). IEEE.

18. Rim, T. H., Lee, C. J., Tham, Y. C., Cheung, N., Yu, M., Lee, G., ... & Wong, T. Y. (2021). Deep-learning-based cardiovascular risk stratification using coronary artery calcium scores predicted from retinal photographs. The Lancet Digital Health, 3(5), e306-e316.

19. Nadakinamani, R. G., Reyana, A., Kautish, S., Vibith, A. S., Gupta, Y., Abdelwahab, S. F., & Mohamed, A. W. (2022). [Retracted] Clinical Data Analysis for Prediction of Cardiovascular Disease Using Machine Learning Techniques. Computational intelligence and neuroscience, 2022(1), 2973324.

20. 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.

21. Reddy, K. V. V., Elamvazuthi, I., Aziz, A. A., Paramasivam, S., Chua, H. N., & Pranavanand, S. (2021). Heart disease risk prediction using machine learning classifiers with attribute evaluators. Applied Sciences, 11(18), 8352.

22. Ramesh, T. R., Lilhore, U. K., Poongodi, M., Simaiya, S., Kaur, A., & Hamdi, M. (2022). Predictive analysis of heart diseases with machine learning approaches. Malaysian Journal of Computer Science, 132-148.

23. Kim, J. O., Jeong, Y. S., Kim, J. H., Lee, J. W., Park, D., & Kim, H. S. (2021). Machine learning-based cardiovascular disease prediction model: a cohort study on the Korean national health insurance service health screening database. Diagnostics, 11(6), 943.

24. Pan, Y., Fu, M., Cheng, B., Tao, X., & Guo, J. (2020). Enhanced Deep learning assisted convolutional neural network for heart disease prediction on the internet of medical things platform. Ieee Access, 8, 189503-189512.

25. Ghosh, P., Azam, S., Jonkman, M., Karim, A., Shamrat, F. J. M., Ignatious, E., ... & De Boer, F. (2021). Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques. IEEE Access, 9, 19304-19326.

26. Ghosh, P., Azam, S., Jonkman, M., Karim, A., Shamrat, F. J. M., Ignatious, E., ... & De Boer, F. (2021). Efficient prediction of cardiovascular disease using machine learning algorithms with relief and LASSO feature selection techniques. IEEE Access, 9, 19304-19326.

27. Jindal, H., Agrawal, S., Khera, R., Jain, R., & Nagrath, P. (2021). Heart disease prediction using machine learning algorithms. In IOP conference series: materials science and engineering (Vol. 1022, No. 1, p. 012072). IOP Publishing.

28. Mathur, P., Srivastava, S., Xu, X., & Mehta, J. L. (2020). Artificial intelligence, machine learning, and cardiovascular disease. Clinical Medicine Insights: Cardiology, 14, 1179546820927404.

29. Yahaya, L., Oye, N. D., & Garba, E. J. (2020). A comprehensive review on heart disease prediction using data mining and machine learning techniques. American Journal of Artificial Intelligence, 4(1), 20-29.

30. Ali, F., El-Sappagh, S., Islam, S. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. S. (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble Deep learning and feature fusion. Information Fusion, 63, 208-222.

31. Predictive modeling for cardiovascular disease risk assessment Christabel T Benila, K Thanammal Jan 2024: i-manager's Journal on Data Science & Big Data Analytics.

32. Industry 4.0 oriented predictive analytics of cardiovascular diseases using machine learning, hyperparameter tuning and ensemble techniques Jameel Ahamed, Mohammad Ahsan Chishti, Roohie Naaz Mir Feb 2022: Industrial Robot: An International Journal

33. Scoping review: Are CVDs risk calculators using the digital platform beneficial for CVDs prevention and management? Mohammed Abd Elfattah Mohammed Darwesh Badawy, Hanif Abdul Rahman, Lin Naing,Nur'Ashikin Tengah, Sofian Johar, Sokking Ong, Nik Ani Afiqah Tuah, Chean Lin Chong Apr 2022

34. Epidemiology and prevention of cardiovascular disease Copy Wenjun Fan, Nathan D Wong, Roger Detels, Alastair H Leyland, Quarraisha Abdool Karim, Fran Baum, Liming Li Nov 2021

35. Cardiovascular disease risk factors prediction using Deep learning convolutional neural networks Mohammad Almatari, Belal Abuhaija, Aladeen Alloubani, Firas Haddad, Ghaith M Jaradat, Yousef Qawqzeh, Mutasem Khalil Alsmadi, Fahad Ali Alghamdi, Jehad Saad Alqurni, Lena Alodat, Linyinxue Dong Aug 2024 : International Journal of Electrical and Computer Engineering (IJECE)

36. Predictive modeling for cardiovascular disease risk assessment Christabel T Benila, K Thanammal Jan 2024 : i-manager's Journal on Data Science & Big Data Analytics

37. Revolutionizing Cardiovascular Health: A Machine Learning Approach for Predictive Analysis and Personalized Intervention in Heart Disease Shilpy Agrawal Mar 2024 : International Journal for Research in Applied Science and Engineering Technology

38. Heart Disease Prediction Using Machine Learning Shivani Yadav Jul 2024 : INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT

39. A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data. Javed Azmi, Md Tabrez Nafis, Guojun Wang, Muhammad Arif, M Afshar Alam, Safdar Tanweer Jul 2022 : Medical Engineering & Physics

40. Early prediction of cardiovascular disease using machine learning: Unveiling risk factors from health records Dr R Deepa, Dr A Sivasamy, Prashant G C, Vijaya Bhaskar Sadu Mar 2024 : AIP Advances

41. Cardiovascular disease risk prediction via machine learning using mental health data M Dorraki, D Abbott, A Van Den Hengel, N Bidargaddi, P J Psaltis, J W Verjans, E Baker, J Narula, Z Liao Dec 2022 : European heart journal. Digital health

42. Enhancing Cardiovascular Disease Risk Prediction with Machine Learning Models Farnoush Shishehbori, Zainab Awan Jan 2024

43. Incorporating longitudinal history of risk factors into atherosclerotic cardiovascular disease risk prediction using deep learning.Jingzhi Yu, Hongyan Ning, Donald M Lloyd-Jones, Lihui Zhao, Xinlei Mi, Lucia C Petito, Xiaoyun Yang, Amy E Krefman, Norrina B Allen, John Wilkins, Lindsay R Pool, Yu Deng Oct 2023 : Research Square.