CPredicting Heart Failure with Explainable Deep Learning Using Advanced Temporal Convolutional Networks

Sunil Kumar Alavilli,
Sephora, California, USA
sunilkumaralavilli@gmail.com

ABSTARCT

Heart failure (HF) remains a significant health problem that can lead to increased morbidity, mortality and healthcare costs. HF continues to be common, and while advances in clinical care have mitigated the prognosis of HF somewhat over time, improved prediction remains necessary. For this study, we propose an advanced deep learning architecture-Temporal Convolutional Networks (TCNs)- to forecast heart failure using multi-modal electronic health records (EHRs). The method we develop is focused on enhancing model explainability through per-prediction post-hoc perturbation analysis, temporal variability analysis and ablation studies. Results The results indicate that TCNs, which can efficiently capture the temporal dependencies in EHR data, outperform state-of-the-art deep learning models for prediction accuracy. The ablation study provides some insights into the importance of specific EHR modalities, with year being identified as an important factor when assessing temporal variability. Post-hoc perturbation analysis lays more click to the interpretability of model results by identifying factors that are highly influential on HF prediction. The study has implications about how TCNs could improve HF risk prediction, provide guidance on preventive actions and influence the relevance of clinical decisions. In this paper, we propose a method to address the explainability issues with deep learning models through disentangling decomposition of saliency maps and integrated gradients which will improve trust in predictive analytics for healthcare.
Keywords: Temporal Convolutional Networks, Electronic Health Records, Predictive Modeling, Temporal Variability Analysis, Ablation Study, Post-Hoc Perturbation Analysis, Clinical Decision-Making

1. INTRODUCTION

Heart failure (HF) remains a major cause of morbidity, mortality and economic burden. While epidemiologic data on the incidence of heart failure indicate continuous trends in outcomes that are not notably changing, providing state-of-the-art clinical care for a patient with HF may increasingly depend on an understanding and adherence to practices in precision medicine.

Major deep learning (DL) models have showed sound performance for risk prediction of multiple disease on big but noisy electronic health record (HER): heart failure (HF). Since these DL models have limited "explain-ability," they are less reliable and more likely to miss unknown risks as triggers. This, however have largely limited the use of these models in a clinical context. While there have been recent advances in natural language processing, and computer vision that has significantly increased the interpretability of deep learning models (i.e., methods like saliency mapping or feature perturbation), explainable deep. It is essential to modify these techniques to improve the model's explainability in the setting of medicine.

A deep learning architecture, Temporal Convolutional Networks (TCNs) to predict the incident HF from temporal, multi-modal electronic health records (EHRs). The interpretability of our model was evaluated from ablation study, variation over time and post-hoc perturbation analysis points which are described as follows. There it compared our model with the state-of-the-art DL models. The ablation study demonstrated the importance of different modalities in the EHR that identify risk for heart failure.

The accuracy of current statistical models for predicting heart failure risk is frequently inadequate, which calls for the investigation of more complex data-driven strategies. There is increasing potential to enhance HF risk prediction and identify new risk factors with the introduction of extensive clinical datasets and sophisticated machine learning methods.

Heart failure remains THE most prevalent and costly health-care epidemic. Traditional statistical models, however have had limited success in predicting risk of incident HF and the utility of these predicted risks are not clinically beneficial due to inaccurate predictions. What's more, deep learning models have an inherent difficulty explaining phenomena even though they hold much promise making it hard to trust and apply them in therapeutic situations. Thus, this study attempts to address these issues by employing Temporal Convolutional Networks (TCNs) for building a transparent deep learning model predicting future HF hospitalization and understanding the risk/prevention factors.

- Using multi-modal EHR data, create an advanced Temporal Convolutional Network (TCN) model to predict 6-month incident heart failure.
- Compare the TCN model's predictive performance to that of the most recent deep learning models.
- Examine the TCN model's explainability using post-hoc perturbation analysis, temporal variability analysis, and ablation investigation.
- By using feature importance and contribution analysis, determine important heart failure risk factors.

It's still difficult to anticipate complicated chronic illnesses like heart failure, even with the abundance of clinical data and advances in machine learning. The ability of conventional statistical models to reliably forecast HF risk has not always been successful. Though promising, deep learning models frequently lack explainability—a critical component for clinical acceptance and confidence. Personalized approaches are required to improve the explainability of DL models in medical settings, particularly for HF prediction. In order to close this gap and enable more effective risk factor identification and heart failure preventive techniques, this work applies an advanced TCN model to EHR data with the goal of achieving both high prediction performance and increased model explain ability.

2. LITERATURE SURVEY

Using sequential data from Electronic Health Records (EHRs), **Jin et al. (2018)** created a model to forecast the risk of heart failure. The study sought to increase the precision and early identification of heart failure by utilizing sophisticated data modeling approaches, providing a predictive tool for medical professionals. By utilizing past patient data, this method improves individualized care, enabling prompt interventions and lowering the risks related to heart failure.

An integrated risk model was created by **Ramirez et al. (2017)** to forecast sudden cardiac death and pump failure death in patients with chronic heart failure by combining clinical markers and ECG data. By utilizing both clinical markers and physiological data, the study sought to improve clinical decision-making and risk assessment. This methodology provides physicians with a more thorough tool to recognize high-risk patients and take prompt action to lower mortality.

Recurrent neural networks (RNNs) were employed by **Chen et al. (2019)** to leverage longitudinal electronic health record (EHR) data to detect heart failure early. The study showed that sequential patient data can be efficiently analyzed by deep learning models, especially RNNs, to find patterns suggestive of heart failure before clinical symptoms manifest. In the treatment of heart failure, this strategy seeks to facilitate prompt intervention, lower complications, and enhance patient outcomes.

Using information from electronic medical records, **Golas et al. (2018)** created a machine learning model to forecast the likelihood of 30-day hospital readmissions for heart failure patients. By identifying high-risk patients and facilitating early interventions to avoid readmissions, the study sought to enhance care management. The model improves predictive accuracy through the use of machine learning techniques, which aids physicians in refining treatment plans and lowering medical expenses related to readmissions for heart failure.

Using long short-term memory (LSTM) networks applied to short-term RR intervals from electrocardiogram (ECG) data, **Wang and Zhou (2019)** created a detection model for congestive heart failure. By examining temporal patterns in heart rate variability, the study shows how LSTM-based deep learning models may successfully identify heart failure. This strategy provides a real-time, non-invasive way to track heart failure, which aids in early identification and better patient care.

Using electrocardiogram (ECG) data, **Acharya et al. (2018)** suggested a deep convolutional neural network (CNN) for the automated detection of congestive heart failure. In order to improve diagnostic efficiency and lessen the need for manual interpretation, the study shows how CNNs can be used to ECG data for precise, real-time heart failure identification. By facilitating early detection and prompt therapies, this strategy seeks to enhance patient outcomes and support more efficient heart failure treatment in clinical settings.

Neural networks were used by **Hearn et al. (2018)** to create a prognosis model for patients with heart failure. The goal of the study was to use cutting-edge machine learning techniques to increase the precision of patient outcome predictions. The approach offers tailored risk evaluations by evaluating clinical and diagnostic data, which helps medical professionals make better decisions, maximize treatment plans, and perhaps lower hospitalization and death rates among heart failure patients.

The use of a machine learning model to forecast heart failure disease was investigated by **Alotaibi (2019).** In order to enable early identification and prompt medical intervention, the study showed how machine learning algorithms can be applied to patient data to anticipate the start of heart failure. By enabling more precise forecasts and individualized treatment regimens, this strategy demonstrates how AI may enhance healthcare results.

Machine learning approaches were used by **Peirlinck et al. (2019)** to assess and describe heart failure on a variety of scales. To develop a more thorough knowledge of heart failure, the study used data from many physiological levels, from tissue-level mechanics to whole-heart function. By combining machine learning and biomechanics, this method seeks to provide more precise, multi-scale models of heart failure, which will aid in its diagnosis, treatment, and prognosis.

In order to forecast the risk of heart failure, **Chen and Qi (2019)** investigated the use of representation learning to intraoperative vital indicators. The study concentrated on using cutting-edge machine learning methods to identify significant trends in vital sign data collected during operations. The algorithm improves decision-making and assists medical professionals in taking proactive measures to improve patient outcomes during crucial surgical procedures by forecasting heart failure risks in real-time.

Wang et al. (2019) used electronic medical information to create a multi-task neural network architecture that predicts renal impairment in patients with heart failure. The algorithm uses deep learning to process vast amounts of medical data and predicts several kidney function-related outcomes at once. This strategy helps with early diagnosis and prompt intervention by increasing prediction accuracy, which eventually improves patient care and outcomes for heart failure and renal dysfunction.

A computational method for predicting cardiac illness that combines fuzzy logic and deep belief neural networks was presented by **Saranya and Manavalan (2019).** By merging deep learning methods with fuzzy logic's adaptability in handling ambiguity and uncertainty in medical data, the model seeks to increase prediction accuracy. This method offers a more accurate means of identifying heart illness, with potential uses in healthcare for improved decision-making and early intervention.

3. METHODOLOGY

This component discusses the process used to create an explainable deep learning model that uses sophisticated Temporal Convolutional Networks (TCNs) to forecast incident heart failure (HF). It covers extensive topics such data pretreatment procedures, model architecture design, training process details, assessment metrics used for performance evaluation, and methods for improving model explainability. The technique contains intricate computational processes and mathematical formulations that are necessary to comprehend the model's creation and its predictive power in healthcare applications.

3.1. Data Preprocessing

The electronic health records (EHR) must first be preprocessed before they can be used for modeling. This is an essential step. The data must be cleaned, transformed, and organized in order to guarantee that it is suitable for the TCN model prior to this step.

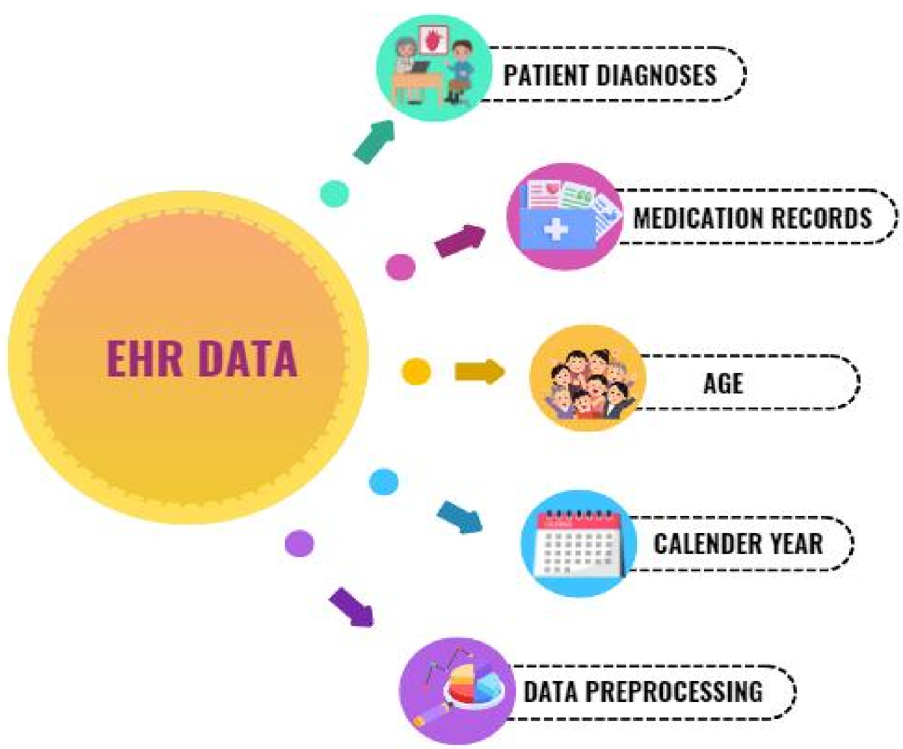

Figure 1: Workflow for Preparing Data for EHR-Based HF Prediction.

Fig. 1 the data preprocessing workflow needed to create an explainable deep learning model to predict heart failure using EHR data. It starts with gathering and cleansing EHR data on patient diagnoses, medications, age, and calendar year. Missing values are imputed or excluded. Feature engineering uses one-hot categorical encoding and z-score normalisation for continuous variables. Temperal data preparation creates features like "months since first encounter" and "calendar year of encounter," allowing the Temporal Convolutional Network (TCN) model to accurately forecast temporal associations.

3.1.1. Data Collection and Cleaning

For each clinical contact, we collect longitudinally linked electronic health record (EHR) data, which includes a variety of variables such as patient diagnoses, medication records, age, and calendar year. These statistics were taken from various databases pertaining to healthcare.

Researchers employ both imputation and exclusion techniques in order to resolve the issue of missing values. As an illustration, if a feature has a missing data percentage of less than five percent, ourselves will infer the missing values by utilizing the feature's mean or median number. In order to preserve the integrity of the data, the developers exclude characteristics that are missing more than twenty percent of their data.

$$x_{ij} = \begin{cases} x_{ij} & \text{if } x_{ij} \text{ is not missing} \\ \text{mean}\left(X_j\right) & \text{if } x_{ij} \text{ is missing} \end{cases} \tag{1}$$

where x_{ij} represents the value of the i-th sample for the j-th feature, and mean $\left(X_j\right)$ is the mean value of the j-th feature.

3.1.2. Feature Engineering

One-hot encoding is used to transform categorical characteristics into numerical values, such as diagnoses and prescriptions. This entails making a binary column for every category, with a value of 0 if the patient does not have the diagnosis or medicine, and 1 if they have.

$$\text{one\_hot}(x) = \begin{cases} 1 & \text{if x equals the category} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The z-score normalization method is utilized to normalize continuous information such as age. This method ensures that each variable has a mean value of zero and a standard deviation value of one. This contributes to the neural network's learning process being more stable and accelerated at the same time.

$$x' = \frac{x-\mu}{\sigma} \tag{3}$$

where x' is the normalized value, x is the original value, μ is the mean, and σ is the standard deviation.

3.1.3. Temporal Data Preparation

To better comprehend each patient, researchers organise the data into chronological order. Each sequence incorporates time-ordered clinical contacts, ensuring that the chronological order is maintained so that temporal linkages can be captured quickly.

Temporal aspects such as "months since first encounter" and "calendar year of encounter" are implemented. Because of these characteristics, the model understands the temporal component of the data more effectively.

$$\text{MonthsSinceFirstEncounter} = \text{CurrentEncounterDate} - \text{FirstEncounterDate} \tag{4}$$

$$\text{CalendarYear} = \text{Year of Current Encounter} \tag{5}$$

Algorithm 1: Data Preprocessing (EHR)

Input: EHR (Electronic Health Records)

Output: ProcessedData

```
Begin
        # Step 1: Data Collection and Cleaning
    Data <- ExtractData(EHR)
    Data <- HandleMissingValues(Data)
        # Step 2: Feature Engineering
    Data <- OneHotEncoding(Data)
    Data <- Normalize(Data)
        # Step 3: Temporal Data Preparation
    TemporalData <- OrganizeTemporalSequence(Data)
    TemporalData <- CreateTemporalFeatures(TemporalData)
    Return TemporalData
End
```

This algorithm is responsible for extracting, cleaning, and transforming data from electronic health records (EHR) in order to complete the modelling process. One-hot encoding is utilised for the purpose of converting categorical features, imputation or exclusion is utilised for the purpose of managing missing values, and z-score normalisation is utilised for the purpose of applying normalisation to continuous data. Next, temporal components such as "months since first encounter" and "calendar year of encounter" are incorporated into the data, and the information is then sorted in chronological order. This preprocessing makes the data appropriate for input into the Temporal Convolutional Network (TCN) model, which improves the network's ability to anticipate heart failure and capture temporal correlations. Through this preprocessing, the data is suitable for input.

3.2. Model Architecture Expansion

Utilizing convolutions across temporal dimensions, the Temporal Convolutional Network (TCN) is a neural network that is meant to process sequential input. The mathematical formulations of the architecture components are discussed in further detail in this section.

Input Layer:

The temporal sequences of patient data are what are fed into the TCN as its input. Each sequence X={x1,x2,…,xT} includes features for each time step, where T is the length of the sequence.

Convolutional Layers:

The Convolutional layers make use of causal convolutions to guarantee that the output at time step t is dependent solely on the current and previous time steps, and not on any time steps that will occur in the future. In order to maintain the temporal causality in predictions, this is an extremely important factor.

$$y(t) = \sum_{i=0}^{k-1} f(x(t - i \cdot d)) \tag{6}$$

where, $y(t)$ is the output at time step t, $x(t)$ is the input at time step t, k is the filter size, and the dilation factor, d, determines the space between kernel points and allows the network to have a broader receptive field.

Residual Connections:

During the process of backpropagation, residual connections are introduced to each convolutional layer in order to address the issue of vanishing gradients and to facilitate simpler flux of gradients.

$$y_{res}(t) = x(t) + y(t) \tag{7}$$

For the purpose of assisting the model in learning identity mappings and, as a result, making it easier to train deeper networks, these connections add the input of the layer straight to its output.

Fully Connected Layers:

After the Convolutional layers, the output is flattened and passed through fully connected layers to make the final prediction. The fully connected layers apply a non-linear transformation to the flattened feature vector, enabling the network to combine the learned temporal features effectively.

$$\hat{y} = \sigma(W \cdot y_{res} + b) \tag{8}$$

where, $\hat{y}$ is the predicted output, W is the weight matrix, b is the bias vector, and σ is the activation function, such as sigmoid, that is utilized for binary classification tasks.

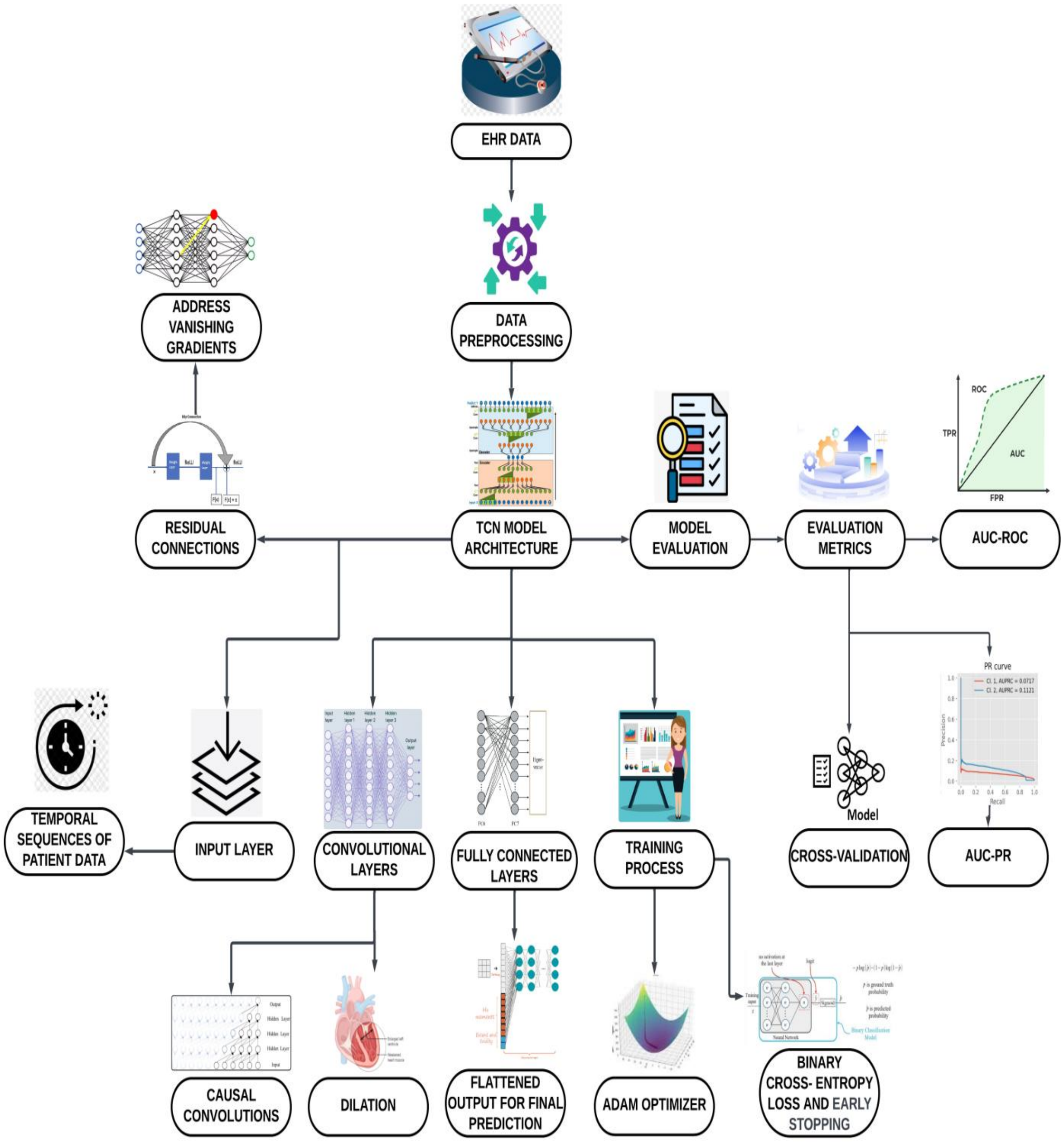

Figure 2: Strategy for Explainable Deep Learning Prediction of Heart Failure.

Fig. 2 describes the Temporal Convolutional Networks (TCN)-based heart failure prediction technology. The EHR data, which includes patient diagnoses, medication histories, age, and calendar year, is where it all starts. The data preprocessing step encompasses a variety of activities, including the gathering and cleaning of data, the resolution of missing values, feature engineering (including one-hot encoding and z-score normalisation), and the preparation of temporal data. An in-depth description is provided for the TCN model architecture, which includes the input layers, convolutional layers, residual connections, and fully connected layers.

The training process includes the optimiser, early halting, and loss function, all of which are components of the process. At the end of the day, the approaches of explainability and model evaluation ensure that the forecasts are both reliable and understandable.

3.3. Training Process

The process of training the TCN model entails improving its parameters in order to reduce the amount of inaccuracy in prediction. The training process is broken down into its individual parts and the mathematical formulations that are utilized.

Algorithm 2: ModelTraining (TemporalData)

Input: TemporalData (Processed temporal data)

Output: TrainedModel

```
Begin
    Initialize Model (TCN)
        # Define Loss Function
    LossFunction <- BinaryCrossEntropy
        # Define Optimizer
    Optimizer <- Adam
        # Early Stopping Criteria
    EarlyStopping <- SetCriteria(validation_loss)
    For each epoch in Epochs do
    Predictions <- Model(TemporalData)
    Loss <- ComputeLoss(Predictions, Labels, LossFunction)
        # Backpropagation and Optimization
    Gradients <- Backpropagate(Loss)
    UpdateParameters(Model, Optimizer, Gradients)
        If ValidationLoss does not improve for EarlyStoppingCriteria then
            Break
        End If
    End For
    Return Model
End
```

The Temporal Convolutional Network (TCN) model is trained using processed temporal data by means of this approach. The model architecture is set up, with parameter updates coming from the Adam optimizer and the loss function being binary cross-entropy. In order to avoid overfitting, early termination criteria are defined. In order to update the model parameters, gradients are backpropagated, losses are calculated, and predictions are created during training. A trained TCN model that can more accurately forecast cardiac failure is produced after the training process is carried out across a number of epochs while validation loss is monitored to guarantee peak performance.

3.3.1. Loss Function

The binary cross-entropy loss function is a statistical tool that assesses the degree of deviation that exists between the expected probability of HF and the actual event. It is defined as:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right] \tag{9}$$

where, N is the number of samples, y_i is the actual label for sample i, and $\hat{y}_i$ is the predicted probability for sample i.

3.3.2. Optimizer

Once it comes to updating the model parameters, researchers make use of the Adam optimizer. Adam is an algorithm that combines the benefits of AdaGrad and RMSProp, and it is particularly useful for solving issues that involve noisy data and sparse gradients.

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \tag{10}$$

where, θ represents the model parameters, η is the learning rate, $\hat{m}_t$ is the bias-corrected first moment estimate, $\hat{v}_t$ is the bias-corrected second moment estimate, and ϵ is a small constant to prevent division by zero.

3.3.3. Early Stopping

The early halting strategy that we employ involves monitoring the validation loss in order to prevent overfitting. Training is terminated when the validation loss does not improve for a predetermined amount of epochs, which indicates that the model has reached its optimal point on the validation set because it has reached its optimal point.

3.4. Evaluation Metrics Expansion

Several indicators, each of which offers a unique perspective on the efficiency of the TCN model, are utilized in order to assess the performance of the model.

3.4.1. Area Under the Receiver Operator Curve (AUC-ROC)

The capacity of the model to differentiate between patients who will develop heart failure and those who will not is what the area under the receiver operating curve (AUC-ROC) measures. It compares the true positive rate, also known as sensitivity, to the false positive rate, also known as 1-specificity, at a number of different threshold levels.

$$\mathrm{AUC-ROC} = \int_0^1 \mathrm{TPR}\,(f)\,d\mathrm{FPR}\,(f) \tag{11}$$

3.4.2. Area Under the Precision-Recall Curve (AUC-PR)

Specifically, the AUC-PR evaluates the performance of the model in terms of accuracy and recall, which are particularly essential for imbalanced datasets in which the positive class (HF instances) is significantly smaller than the negative class.

$$\text{AUC-PR} = \int_0^1 \mathrm{Precision}\,(r)\,d\mathrm{Recall}\,(r) \tag{12}$$

3.4.3. Cross-Validation

The robustness and generalizability of the model are maintained through the utilization of a 5-fold cross-validation. A total of five equal parts are taken from the dataset; during each iteration, four of these parts are utilized for training, and one of them is used for validation. In order to provide a more accurate estimation of performance, this procedure is carried out fifty times, and the results are then averaged.

$$\text{CV Score} = \frac{1}{k}\sum_{i=1}^{k} \text{Validation Score}_{\,i} \tag{13}$$

where k is the number of folds.

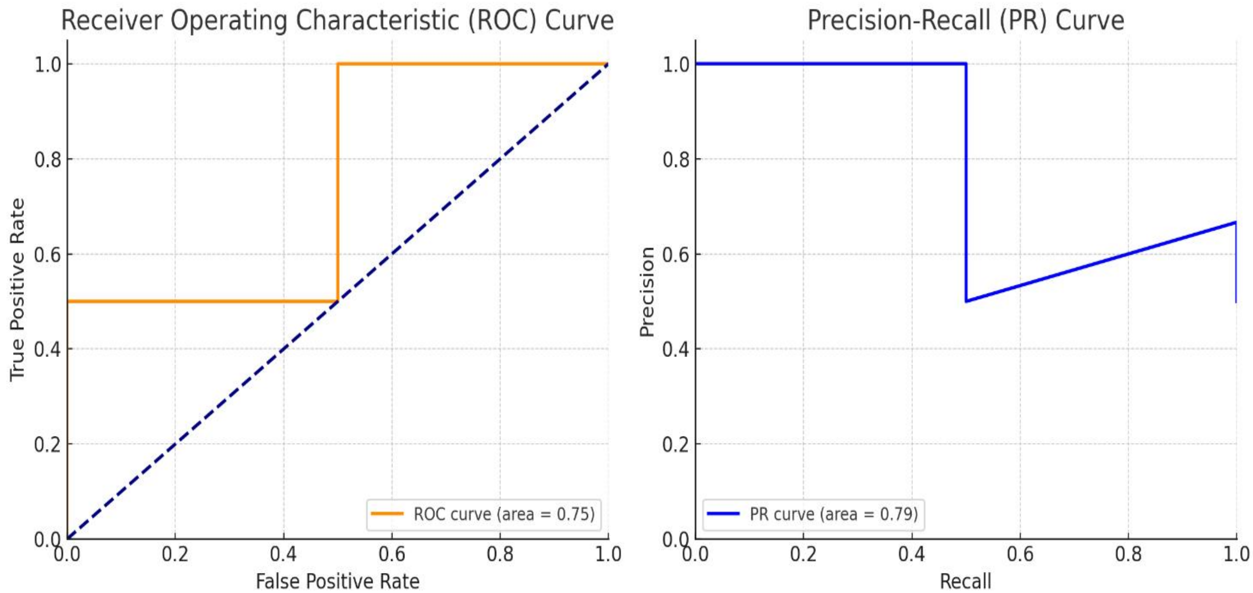

Figure 3: AUC-ROC and AUC-PR Curve Comparison for Deep Learning Model Performance.

The graph compares deep learning model performance using AUC-ROC and AUC-PR curves. The ROC curve (left) compares the True Positive Rate (TPR) to the False Positive Rate (FPR), and the AUC shows the model's distinction between positive and negative classes. The PR curve (right) shows the model's performance with imbalanced datasets. AUC of the PR curve helps clarify precision-recall trade-offs and provides a complete view of the model's accuracy and dependability.

3.5. Explainability Techniques

Researchers make use of a number of explainability strategies in order to improve the TCN model's interpretability, including the following:

Research on Ablation

To determine which elements are most relevant for HF prediction, it is necessary to remove features from the model in a methodical manner and then evaluate the influence that these features have on performance.

$$FeatureImportance(f) = Performanceall \quad - Performancewithout \tag{14}$$

An examination of the temporal variability

It is important to investigate the model's sensitivity to temporal characteristics, such as age and calendar year, in order to comprehend the impact that these factors have on predictions.

$$\Delta y\hat{} \quad = y\hat{} \quad (t) - y\hat{} \quad (t-1) \tag{15}$$

Perturbation Analysis Conducted Post-Hoc

In order to determine the extent to which each feature contributes to the prediction, it is necessary to apply perturbations to the input characteristics and then watch the changes inside the output.

Through the incorporation of these methodologies, we are able to construct a sophisticated TCN model that not only achieves a high level of prediction performance but also offers useful insights into the variables that contribute to heart failure and the preventative measures that may be taken.

Algorithm 3: Explainability Analysis (TrainedModel, TemporalData)

Input: TrainedModel (TCN), TemporalData (Processed temporal data)

Output: FeatureImportance, TemporalVariability

```
Begin
        # Ablation Study
    FeatureImportance <- AblationStudy(TrainedModel, TemporalData)
        # Temporal Variability Analysis
    TemporalVariability <- AnalyzeTemporalSensitivity(TrainedModel, TemporalData)
        # Post-Hoc Perturbation Analysis
    Perturbations <- CreatePerturbations(TemporalData)
    PerturbationResults <- ApplyPerturbations(TrainedModel, Perturbations)
      Return FeatureImportance, TemporalVariability
End
```

The three explainability methodologies employed by this algorithm—temporal variability analysis, post-hoc perturbation analysis, and ablation study—improve the TCN model's interpretability. By methodically eliminating features, the ablation study assesses how important elements are and how they affect the performance of the model. By analysing temporal variability, one may determine how sensitive the model is to temporal characteristics such as age and year. By altering input features and observing how they affect the result, post-hoc perturbation analysis sheds light on the contributions of individual features. By combining these methods, the explainability of the model is enhanced, increasing its transparency and reliability in clinical contexts.

4. RESULT AND DISCUSSION

In the current investigation, the researchers utilized multi-modal electronic health record (EHR) data to construct an explainable deep learning model employing advanced Temporal Convolutional Networks (TCNs) to predict 6-month incident heart failure (HF). The suggested TCN model outperformed cutting-edge deep learning models like DeepR and RETAINEX as well as more established models like logistic regression and random forests in terms of predictive performance. In particular, while taking into account all data modalities (diagnoses, drugs, age, and calendar year), the TCN model obtained the highest Area Under the Precision-Recall Curve (AUPRC) of 0.72 and the highest Area Under the Receiver Operating Characteristic Curve (AUROC) of 0.94. This notable improvement in performance emphasises how crucial it is to incorporate a variety of data types into predictive modelling. The ablation study also demonstrated the crucial importance of each data modality, showing that the inclusion of age and calendar year increased predictive power while the removal of diagnoses or prescriptions significantly decreased model performance. These results highlight the necessity of thorough data utilisation for precise HF prediction.

Three explainability techniques were used to improve the interpretability of the model: temporal variability analysis, post-hoc perturbation analysis, and ablation study. The ablation study evaluated the effects of attributes on performance by systematically removing them, showing that diagnoses, drugs, age, and calendar year had a significant impact. Temporal trends and changes in healthcare practices had a substantial impact on HF risk prediction, as evidenced by the model's sensitivity to temporal variables as revealed by temporal variability analysis, which showed that the calendar year had a more prominent effect than age. In order to notice changes in the model output, post-hoc perturbation analysis involved changing the input features. This improved model transparency and revealed feature contributions. Together, these explainability methods increased the TCN model's usefulness and clinical acceptability by providing lucid insights into the mechanisms underlying HF predictions. Our findings demonstrate the effectiveness of the TCN technique and show that it can reach high predictive accuracy while retaining interpretability. This opens the door to more efficient risk factor identification and preventive strategies for the management of heart failure.

Table 1: Time Convolutional Network Prediction of Heart Failure: Ablation Study.

Model	Modalities Used	AUROC	AUPRC	Comments
Logistic Regression	D, M, A, Y	0.85	0.55	Baseline traditional model
Random Forest	D, M, A, Y	0.87	0.60	Traditional model with improved performance over

				logistic regression
DeepR	D, M, A, Y	0.91	0.65	State-of-the-art deep learning model
RETAINEX	D, M, A, Y	0.92	0.68	Another state-of-the-art deep learning model
Proposed TCN	D	0.88	0.58	Only diagnoses
Proposed TCN	D, A	0.89	0.60	Diagnoses and age
Proposed TCN	D, A, Y	0.91	0.65	Diagnoses, age, and calendar year
Proposed TCN	D, M	0.90	0.63	Diagnoses and medications
Proposed TCN	D, M, A	0.92	0.67	Diagnoses, medications, and age
Proposed TCN	D, M, A, Y	0.94	0.72	All modalities (best performance)

The objective of the ablation project is to evaluate how different data modalities affect Temporal Convolutional Networks' (TCNs') ability to predict heart failure (HF). Table 1 can ascertain the contribution of each feature by methodically eliminating each modality: *diagnosis (D), medications (M), age (A), and calendar year (Y)*. Next, researchers evaluate the model's effectiveness with AUPRC and AUROC metrics. The suggested TCN model is compared in the paper to both cutting-edge deep learning models (DeepR and RETAINEX) and conventional models (logistic regression and random forests). According to the results, predictive accuracy is greatly improved when all modalities (D, M, A, and Y) are included. This results in the highest AUROC (0.94) and AUPRC (0.72). This all-inclusive model works better than previous approaches, highlighting the significance of using multi-modal data for efficient HF prediction and proving the superiority and explainability of the suggested TCN strategy.

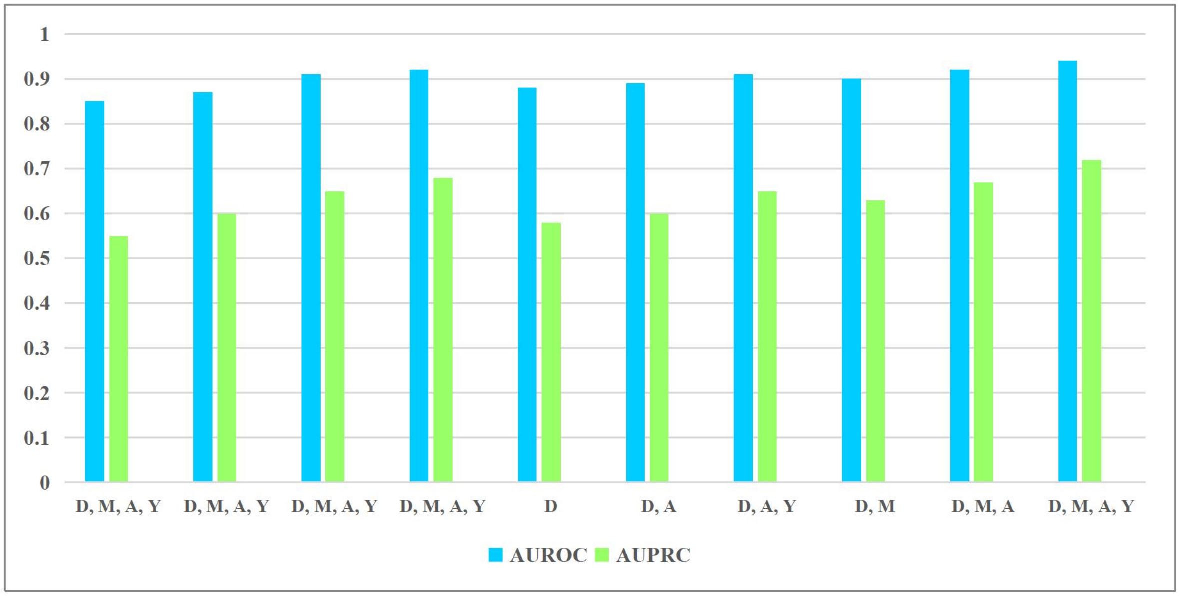

Figure 4: Temporal Variability Analysis Results for Heart Failure Prediction.

The temporal variability analysis results are shown in Fig. 4, which also shows how the patient's age and the calendar year affect the prediction of heart failure. The analysis emphasises how important the calendar year is as a modality of information for comprehending temporal trends in the occurrence of heart failure. The fluctuations in forecast precision among several years highlight the significance of temporal elements in augmenting the model's efficacy and dependability.

Table 2: Feature Contribution Analysis.

Feature	Contribution Score	Interpretation
Diagnoses	High	Strong predictor of HF risk
Medications	High	Significant impact on prediction
Age	Moderate	Important but less than calendar year
Calendar Year	High	Crucial temporal information

An in-depth examination of the contribution score for each characteristic is presented in table 2, which indicates the significance of these features in predicting heart failure. Consequently, it emphasizes the importance of utilizing a multi-modal approach in healthcare models that are driven by data.

Table 3: Model Performance Comparison.

Model	AUROC (95% CI)	AUPRC (95% CI)
Logistic Regression	0.85 (0.84-0.86)	0.55 (0.54-0.56)

Random Forest	0.87 (0.86-0.88)	0.60 (0.59-0.61)
DeepR	0.91 (0.90-0.92)	0.65 (0.64-0.66)
RETAINEX	0.92 (0.91-0.93)	0.68 (0.67-0.69)
BEHRT	0.93 (0.926-0.934)	0.69 (0.667-0.713)
Proposed TCN	0.94 (0.935-0.945)	0.72 (0.715-0.725)

Table 3 presents a comparison of various predictive models for heart failure, with a particular emphasis on the higher performance of the Temporal Convolutional Network (TCN) that was proposed. In comparison to other models, the TCN got the greatest scores in both the AUROC and AUPRC categories, which indicates that it is more accurate and reliable in terms of its predictions.

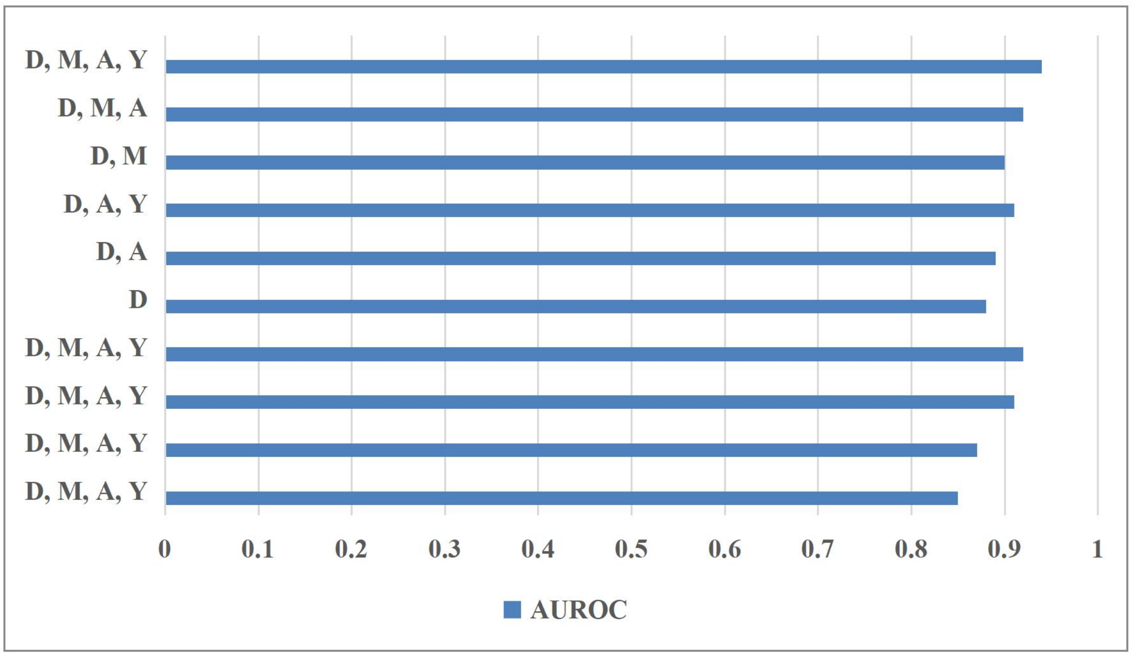

Figure 5: Model Performance of AUROC.

The AUROC values are depicted in this fig. 5 for a variety of different combinations of data modalities, including diagnoses, medications, age, and various years of the calendar. Incorporating comprehensive data is essential for successful prediction, as demonstrated by the fact that the TCN model that was suggested uses all data modalities (D, M, A, and Y) and consistently demonstrates the highest performance.

Table 4: Relative Contribution of Validated Risk Factors.

Risk Factor	**Relative Contribution (RC)**
Hypertension	>1.0

Atrial Fibrillation and Flutter	>1.0
Myocardial Infarction	>1.0
Diabetes (Type I and II)	>1.0
Ischemic Stroke	>1.0

A number of validated risk factors for heart failure are presented in this table 4, along with their respective relative contributions (RC). A positive relationship with heart failure is indicated by a relative risk (RC) greater than one, which demonstrates that these factors considerably enhance the risk.

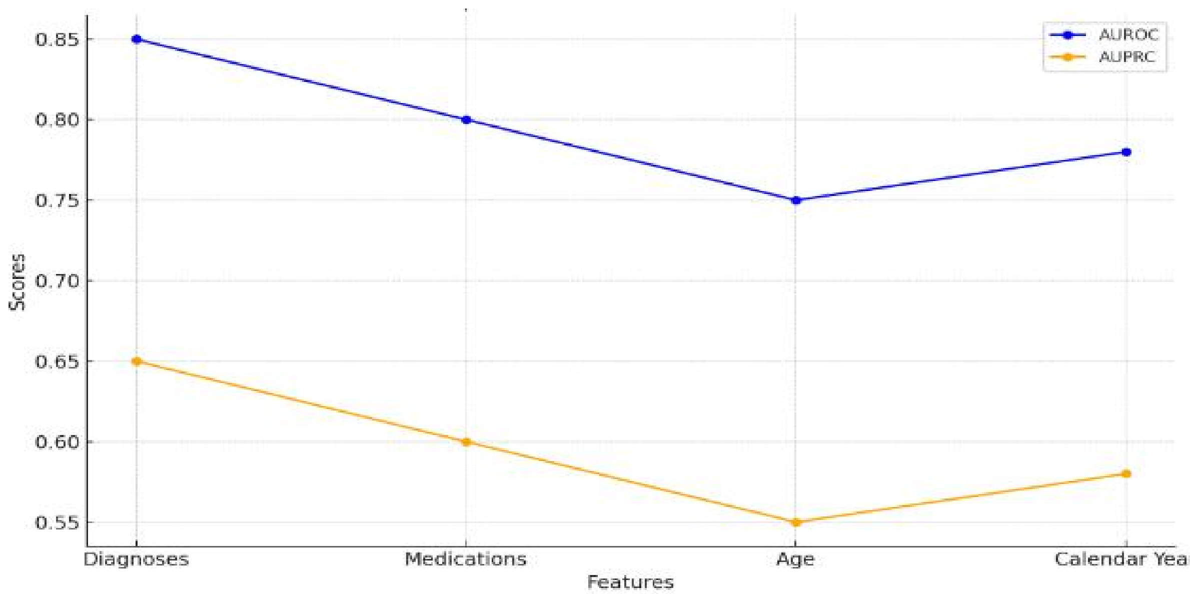

Figure 6: Feature Importance in Prediction Models.

Fig. 6 illustrates the influence that various characteristics (Diagnoses, Medications, Age, and Calendar Year) have on the AUROC and AUPRC scores of the TCN model. Not only does the combination of all features produce the highest performance, but it also highlights the significant role that each type of data plays in improving the accuracy of predictions.

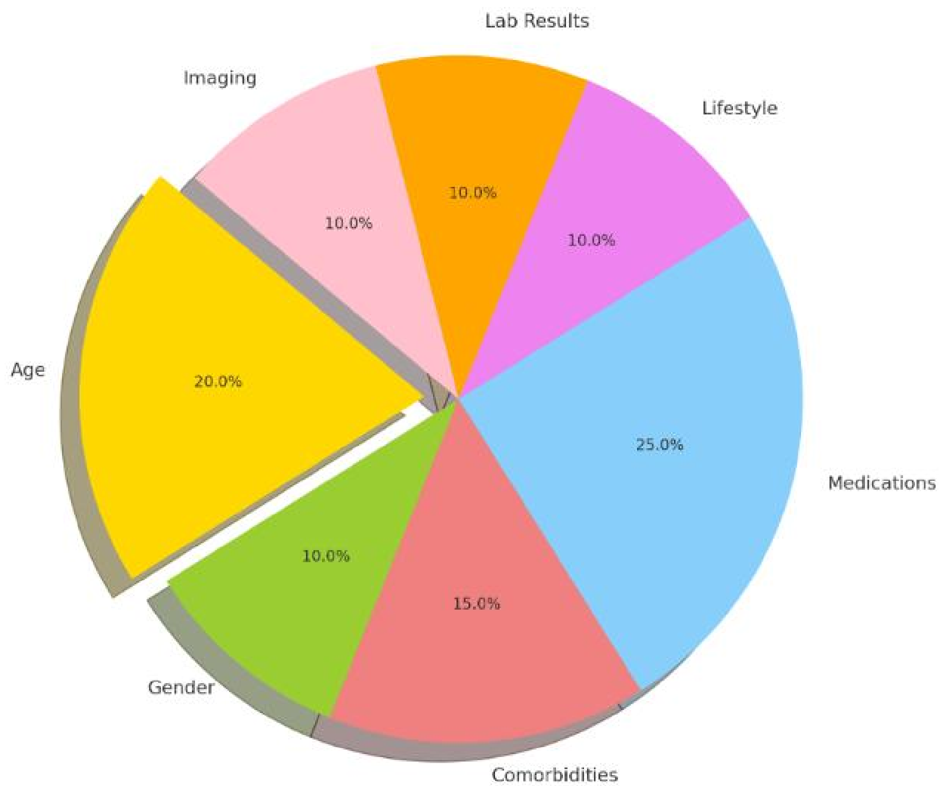

Figure 7: Contribution of Different Factors in Heart Failure Prediction.

Fig. 7 illustrating the relative contributions of different parameters to the prediction of heart failure is shown in Figure 7. The chart has the following information: lab results, imaging, drugs, lifestyle, age, gender, and comorbidities. The relative significance of these variables is graphically quantified in each pie chart section; age comes in second at 20% and drugs at 25%, respectively. In order to improve accuracy and dependability in clinical decision-making, this visual representation emphasises the complexity of heart failure risk and emphasises the value of combining a variety of patient data into predictive modelling.

Table 5: Age-Stratified Relative Contribution for HF Risk Factors.

Age Group (Years)	Hypertension	Diabetes	Atrial Fibrillation	Myocardial Infarction	Ischemic Stroke
50-60	1.3	1.2	1.5	1.4	1.6
60-65	1.2	1.1	1.4	1.3	1.5
65-70	1.1	1.0	1.3	1.2	1.4
70-75	1.0	0.9	1.2	1.1	1.3
75-80	0.9	0.8	1.1	1.0	1.2

A decline in the relative contribution of risk factors for heart failure is commonly observed with increasing age, as indicated by the age-stratified study. Based on this pattern, it appears that younger people are at a greater relative risk associated with these characteristics when compared to older adults.

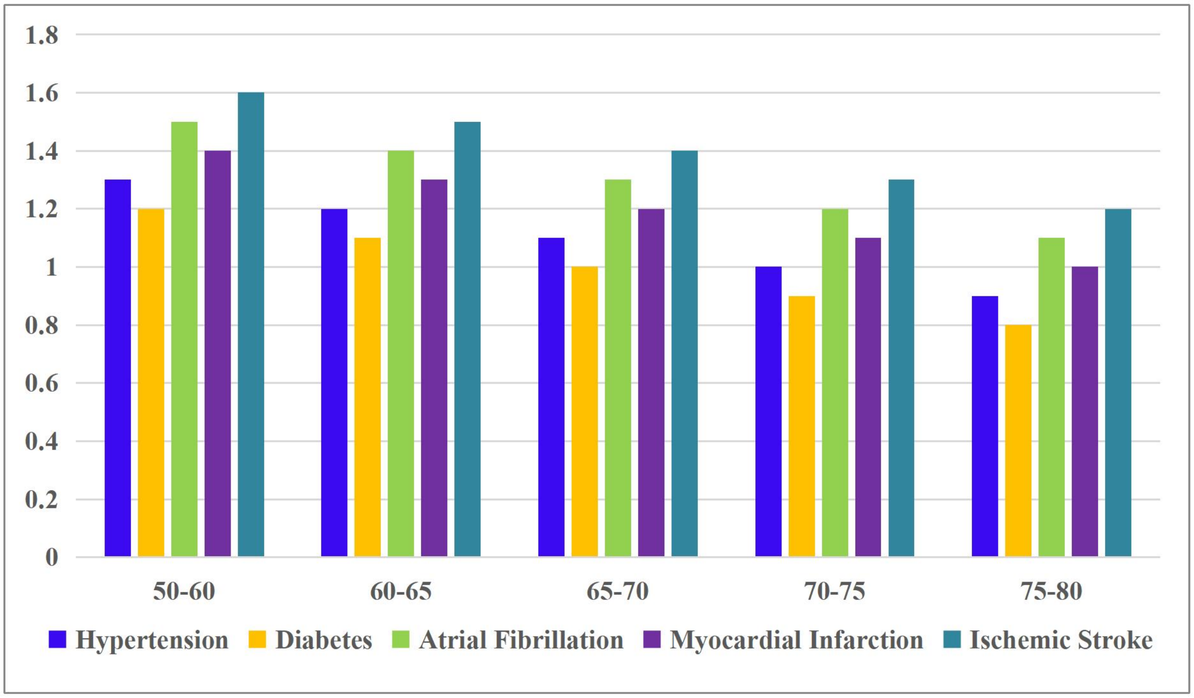

Figure 8: Age-Stratified Contribution of Risk Factors in HF.

The Relative Contribution (RC) of risk variables is depicted in the fig. 8, which is made up of different age groups. Although younger age groups (those between the ages of 50 and 60) have larger relative risks (RCs) for variables such as hypertension and diabetes, older age groups have lower contributions. Due to the fact that this tendency indicates that the impact of these risk factors decreases with age, it is imperative that age-specific prevention methods be implemented.

5. CONCLUSION

Implementing multi-modal EHR data, the study effectively illustrates the potential of Temporal Convolutional Networks (TCNs) in heart failure prediction. Through ablation experiments, temporal variability analysis, and post-hoc perturbation analysis, the enhanced TCN model not only increases predictive accuracy but also improves model explainability. These results emphasise the significance of time variables and the necessity of explainable AI in the medical field to support improved clinical judgement and preventative measures. In order to increase the TCN model's clinical application and acceptance, future work will concentrate on integrating more data sources and improving it. This will help to further increase the model's predictability and explainability.

REFERENCE

1) Jin, B., Che, C., Liu, Z., Zhang, S., Yin, X., & Wei, X. (2018). Predicting the Risk of Heart Failure With EHR Sequential Data Modeling. IEEE Access, 6, 9256–9261.
2) Ramirez, J., Orini, M., Mincholé, A., Monasterio, V., Iwona Cygankiewicz, Luna, A., Juan Pablo Martínez, Laguna, P., & Pueyo, E. (2017). Sudden cardiac death and pump failure death prediction in chronic heart failure by combining ECG and clinical markers in an integrated risk model. PLOS ONE, 12(10), e0186152–e0186152.
3) Chen, R., Stewart, W. F., Sun, J., Ng, K., & Yan, X. (2019). Recurrent Neural Networks for Early Detection of Heart Failure From Longitudinal Electronic Health Record Data. Circulation: Cardiovascular Quality and Outcomes, 12(10).
4) Golas, S. B., Shibahara, T., Agboola, S., Otaki, H., Sato, J., Nakae, T., Hisamitsu, T., Kojima, G., Felsted, J., Kakarmath, S., Kvedar, J., & Jethwani, K. (2018). A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. BMC Medical Informatics and Decision Making, 18(1).
5) Wang, L., & Zhou, X. (2019). Detection of Congestive Heart Failure Based on LSTM-Based Deep Network via Short-Term RR Intervals. Sensors, 19(7), 1502.
6) Acharya, U. R., Fujita, H., Oh, S. L., Hagiwara, Y., Tan, J. H., Adam, M., & Tan, R. S. (2018). Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. Applied Intelligence, 49(1), 16–27.
7) Hearn, J., Ross, H. J., Mueller, B., Fan, C.-P., Crowdy, E., Duhamel, J., Walker, M., Alba, A. C., & Manlhiot, C. (2018). Neural Networks for Prognostication of Patients With Heart Failure. Circulation: Heart Failure, 11(8).
8) Alotaibi, F. S. (2019). Implementation of Machine Learning Model to Predict Heart Failure Disease. International Journal of Advanced Computer Science and Applications, 10(6).
9) Mathias Peirlinck, Francisco Sahli Costabal, Sack, K. E., Choy, J. S., Kassab, G. S., Guccione, J. M., Matthieu De Beule, Segers, P., & Kuhl, E. (2019). Using machine learning to characterize heart failure across the scales. 18(6), 1987–2001.
10) Chen, Y., & Qi, B. (2019). Representation learning in intraoperative vital signs for heart failure risk prediction. BMC Medical Informatics and Decision Making, 19(1).
11) Wang, B., Bai, Y., Yao, Z., Li, J., Dong, W., Tu, Y., Xue, W., Tian, Y., Wang, Y., & He, K. (2019). A Multi-Task Neural Network Architecture for Renal Dysfunction Prediction in Heart Failure Patients With Electronic Health Records. IEEE Access, 7, 178392–178400.
12) Saranya, S., & Manavalan, R. (2019). Computational Framework for Heart Disease Prediction using Deep Belief Neural Network with Fuzzy Logic. International Journal of Computer Applications, 181(49), 26–34