Seminars in Medical Writing and Education. 2023; 2:126
doi: 10.56294/mw2023126

ORIGINAL

Integrating Clinical Data and Advanced Analytics for Accurate Readmission Predictions in Heart Failure

Integración de datos clínicos y análisis avanzados para predecir con exactitud los reingresos por insuficiencia cardíaca

Deepak Kumar Parhi[1], Pooja Varma[2], Supriya Awasthi[3]

[1]IMS and SUM Hospital, Siksha 'O' Anusandhan (deemed to be University), Department of Cardiology. Bhubaneswar, Odisha, India.
[2]JAIN (Deemed-to-be University), Department of Psychology. Bangalore, Karnataka, India.
[3]School of Allied Health Sciences, Noida International University. Greater Noida, Uttar Pradesh, India.

Cite as: Parhi DK, Varma P, Awasthi S. Integrating Clinical Data and Advanced Analytics for Accurate Readmission Predictions in Heart Failure. Seminars in Medical Writing and Education. 2023; 2:126. https://doi.org/10.56294/mw2023126

Submitted: 11-09-2022 **Revised:** 23-12-2022 **Accepted:** 27-02-2023 **Published:** 28-02-2023

Editor: PhD. Prof. Estela Morales Peralta

Corresponding Author: Deepak Kumar Parhi

ABSTRACT

Heart failure (HF) is a chronic and progressive condition that significantly impacts healthcare systems worldwide due to high hospitalization and readmission rates. Traditional prediction models frequently rely on clinical assessments and historical data, but they cannot provide the accuracy required for effective intervention. Integrating clinical data with advanced analytics offers a promising approach to improving readmission prediction models, enabling targeted interventions for high-risk patients. The research aimed to develop an accurate 30 days and 60 days readmission prediction model for HF patients using clinical data and deep learning (DL) techniques. An Efficient Cockroach Swarm-tuned Deep Belief Network (ECS-DBN) model is provided to predict the risk of readmission in heart failure patients. The dataset included HF clinical information in readmissions. The dataset included health records, diagnostic test results, treatment history, and patient demographics. Data cleaning and normalization are performed to ensure accuracy and consistency. Efficient cockroach swarm optimization is employed to fine-tune the hyperparameters of the DBN, enhancing its predictive accuracy and computational efficiency of readmission in heart failure patients. An ideal categorization threshold was established based on anticipated cost reductions, and performance was evaluated using the correlation statistic. The ECS-DBN model outperformed other techniques, achieving a high accuracy (0,96), recall (0,63), precision (0,97), F1-Score (0,65) and recall (0,63) compared to the conventional method in 60 days. The results show that using advanced analytics to analyze clinical data enhances the prediction of readmission in patients with HF. By identifying high-risk individuals early on, the suggested paradigm optimizes healthcare by enabling focused interventions and improving clinical outcomes.

Keywords: Heart Failure (HF); Readmission Predictions; Clinical Data; an Efficient Cockroach Swarm-Tuned Deep Belief Network (ECS-DBN).

RESUMEN

La insuficiencia cardiaca (IC) es una enfermedad crónica y progresiva que afecta significativamente a los sistemas sanitarios de todo el mundo debido a las elevadas tasas de hospitalización y reingreso. Los modelos de predicción tradicionales suelen basarse en evaluaciones clínicas y datos históricos, pero no pueden proporcionar la precisión necesaria para una intervención eficaz. La integración de datos clínicos con análisis avanzados ofrece un enfoque prometedor para mejorar los modelos de predicción de reingresos, lo que permite intervenciones específicas para pacientes de alto riesgo. La investigación tuvo como objetivo desarrollar un modelo preciso de predicción de readmisión a 30 y 60 días para pacientes con IC utilizando datos clínicos y técnicas de aprendizaje profundo (DL). Se proporciona un modelo de Red de Creencias Profundas (ECS-DBN, por sus siglas en inglés) ajustado por enjambre de cucarachas para predecir el riesgo de reingreso en pacientes con insuficiencia cardiaca. El conjunto de datos incluía información clínica de IC en reingresos. El conjunto de datos incluía historiales médicos, resultados de pruebas diagnósticas, historial de tratamiento y datos demográficos del paciente. Se realiza una limpieza y normalización de los datos para garantizar su precisión y coherencia. Se emplea la optimización eficiente de enjambre de cucarachas para ajustar los hiperparámetros de la DBN, mejorando su precisión predictiva y eficiencia computacional de readmisión en pacientes con insuficiencia cardiaca. Se estableció un umbral de categorización ideal basado en la reducción de costes prevista y se evaluó el rendimiento mediante el estadístico de correlación. El modelo ECS-DBN superó a otras técnicas, logrando una alta exactitud (0,96), recuperación (0,63), precisión (0,97), F1-Score (0,65) y recuperación (0,63) en comparación con el método convencional en 60 días. Los resultados muestran que el uso de analítica avanzada para analizar datos clínicos mejora la predicción de reingresos en pacientes con IC. Al identificar precozmente a los individuos de alto riesgo, el paradigma sugerido optimiza la asistencia sanitaria al permitir intervenciones focalizadas y mejorar los resultados clínicos.

Palabras clave: Insuficiencia Cardiaca (IC); Predicciones de Readmisión; Datos Clínicos; una Red de Creencia Profunda Eficiente Sintonizada por Enjambre de Cucarachas (ECS-DBN).

INTRODUCTION

Heart failure (HF) is a heterogeneous disease; it was initially identified as a developing crisis about 25 years ago. The entire quantity of coronary illness patients is continually rising due to an aging and increasing population.[1] The causes of hospital readmissions in HF are complex and have multiple concerns, including disease-centered factors, such as worsening heart activities, and healthcare system factors, like inadequate care.[2] The primary cause of hospital readmissions in the real-world community of HF patients frequently happens in their earlier post-discharge period. These risk variables differ depending on the reason and timing of hospital readmissions.[3] In individuals with HF, concurrent conditions, like diabetes, high blood pressure, anemia, atrial fibrillation, diuretics, renal failures, and increased thyroid were linked to greater rates of readmission.[4] To prevent readmission of people, raise the standard of care, lower healthcare system expenses, and address the patient needs for improved care, a suitable prediction and analytics system is required.[5] The convention HF prediction methods rely on statistical models and clinical scoring; the complexity of handling large information demands cutting-edge methods, like artificial intelligence (AI) to improve timely diagnosis, risk evaluation, and tailored therapies.[6] With the application of new technologies, the AI-driven models can overcome the limitations of the traditional approaches by enhancing the prediction accuracy, and overall patient outcomes and reducing HF patient readmission.[7]

The research evaluated the deep learning (DL) architectures to forecast the 30-day critical care readmission probability caused by a variety of diseases.[8] For the data gathering purpose, the MIMIC-III data was employed. Multiple DL methods were used, such as neural ordinary differential equations (ODEs), recurrent layers, systems of attention, and medical concept embeddings. For static variables, odds ratios and subsequent weights were calculated using Bayesian estimation. The outcomes showed that a recurrent neural network (RNN) produced the best F1-score (0,372), AUROC (0,739), and accuracy (0,331). Models based on attention provided accessibility with minimal degradation of accuracy.

To increase quick readmission prediction and identify concerns, like class imbalance and missing data, the research utilized machine learning (ML) techniques on patient information.[9] The information form 1856 HF patient details were gathered, which included laboratory tests, health indicators, hospitalizations historical events, and characteristics. The six ML models, including support vector machine (SVM), naïve bayes (NB), least-square SVM (LS-SVM), bagging, random forest (RF) and AdaBoost, were assessed. The findings showed that RF obtained the scores of accuracy (0,91) demonstrated the greater performance levels.

To overcome disparities in class concerns, the investigation implemented a machine learning (ML) model that could predict 30-day HF admission.[10] The examination obtained the necessary information from the 10757 HF patients in 7 years. A multi-layer perceptron (MLP) based model was implemented with necessary preprocessing and feature extraction. From the findings, it was concluded that MLP outperformed other traditional techniques in terms of an AUC of 0,62, specificity of 70,01, and sensitivity of 48,42 leading to improvements in the accuracy of HF readmission and mortality prediction and having the best predictive performance.

The investigation employed different ML algorithms to predict HF patients in the re-hospitalization process within the specific duration.[11] The information was acquired under the guidance of a skilled cardiologist for the reliability and the key mechanisms of extracting features and the pre-processing were performed. The different algorithms of logistic regression (LR), SVM, decision trees (DT), and artificial neural networks (ANN) were implemented. The findings showed the changing relative relevance over time, offered guidance for high-quality, reasonably priced HF patient care, and lowered readmission rates.

The evaluation used ML techniques to predict hospital readmissions for individuals with atrial fibrillation in the HF.[12] The 2013 nationwide readmissions source was used to collect data, with a particular focus on 30-day readmissions. The ML models comprised of DT and SVM were applied. In addition, the k-nearest neighbors (KNN) also employed. KNN delivered superior accuracy (0,84), specificity (0,997), precision (0,886), AUC (0,91), and sensitivity (0,713), which was higher than the two ML approaches. These results proved the effective risk assessment and preventative measures by highlighting important factors driving HF problems.

To forecast hospital readmissions among patients with HF and address missing data in electronic health records (EHR) the research introduced a novel method.[13] The MIMIC-III database was employed as the data collection process. For the regeneration of data that was absent, the Gaussian Process Latent Variable Model (GPLVM), was performed, which developed a lower-dimensional encoding to predict values. A constrained support vector machine (cSVM) was suggested for obtaining characteristics, taking input uncertainty into account for reliable forecasting. With an AUC of 0,68, the results revealed a 7 % increase in accuracy in predicting and adjusted mean absolute mistakes of 0,11–0,12 compared to traditional methods.

The examination provided a framework for quickly and accurately predicting HF mortality.[14] The number of 10,198 inpatient data from the regional hospital was utilized for data collection. A DL method of feature rearrangement-based deep learning system (FRDLS) was recommended and the imbalance in data is addressed and is also employed as feature extraction. The outcomes revealed that the model predicts in-hospital, 30-day, and 1-year mortality with excellent accuracy and AUC (90,37), outperforming conventional ML techniques.

The investigation developed a neural network rather than logistic regression to predict 30 days of rehospitalization in people with HF.[15] The administration asserts that they were used to gather information on 343,328 HF hospitalizations and a stratified 5-fold cross-validation technique is employed. Among the AI models, the LR did well at 0,643 AUC, while an RNN-CRF model obtained 0,642 AUC.[16] The findings demonstrate that hospitalization schedules enhance forecasts and that administrative data outperforms clinical information in terms of competitiveness.

The purpose of the research is to develop an efficient ESC-DBN model for improving the accuracy and reliability of hospital readmission prediction in HF individuals that contributes to reducing unnecessary readmissions, improving patient care, and optimizing hospital resource allocation.

The remaining part of the research is organized as follows: Section 2 describes the implemented ECS-DBN model with necessary data gathering and preprocessing approaches. Section 3 concentrates on the effective findings obtained from the investigation. Section 4 offers a detailed discussion of the overall outcomes and the conclusion is narrated in the final Section of Part 5.

METHOD

A brief explanation of the implemented ECS-DBN model with effective visualizations, the major role of the preprocessing mechanism, which concludes the data cleaning, normalization, and feature extraction approaches is all delivered in the part. Figure 1 displays the outline flow of the suggested ECS-DBN model.

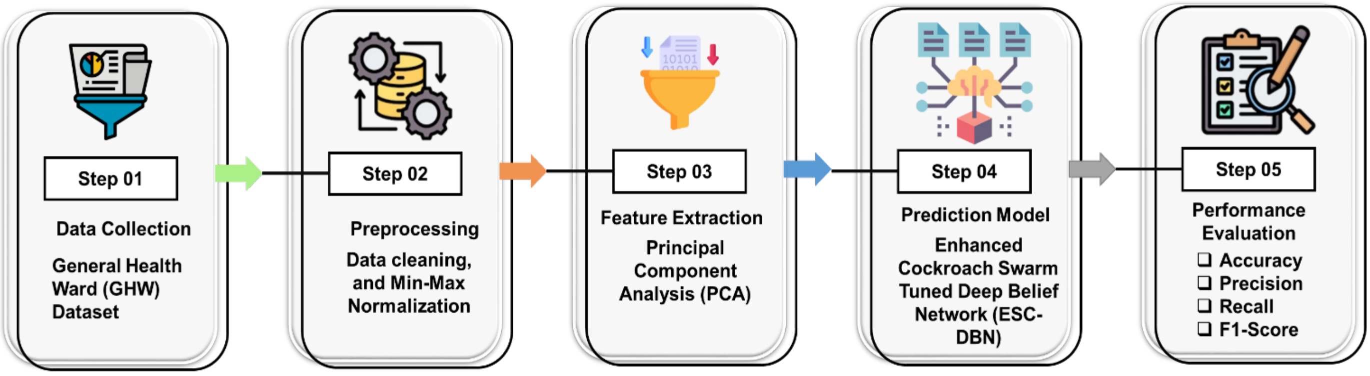

Figure 1. ECS-DBN model flow

Data collection

The General Hospital Ward (GHW) dataset is utilized for HF readmission analysis, specifically predicting 30- and 60-day hospital readmissions for coronary artery disease patients.[17] The crucial factors such as health records, diagnostic test results, treatment history, patient demographics, and HF-specific characteristics are taken into account. The model accuracy is improved by integrating these diverse clinical variables for enhanced prediction.

Data preprocessing

The data pre-processing transforms raw clinical data into a standardized and organized format for DL-based readmission prediction in HF. Two important mechanisms are involved in the stage, which is explained below.

Data cleaning

Initially, when dealing with abnormalities and missing clinical variables, the data cleaning process minimizes the irrelevant data that does not aid in model construction. The looping records were identified and removed to prevent data redundancy. Additionally, categorical variables, such as comorbidities and medication history, were standardized. Continuous variables undergo normalization to ensure consistency in scale, which enhances model convergence.

Normalization

The process of normalizing the clinical features is achieved through the Min-Max normalization. Normalization played a crucial role in ensuring the constant feature ranges for enhanced model performance by scaling the clinical variables. The process of preventing high-magnitude characteristics from dominating while maintaining differences in relative magnitude is achieved by transforming values to a predefined range of 0 to 1. Equation 1 demonstrates the normalizing formula for each feature.

$$y_{new} = \frac{y - \min(y)}{\max(y) - \min(y)} \qquad (1)$$

Where the compensated value that is computed from the normalized outcomes is denoted by y_{new}. max(y) and min(y) are the highest and lowest values in the dataset. For enhancing convergence speed and stability in the ECS-DBN model, the Min-Max normalization is essential.

Feature extraction

The principal component analysis (PCA) is one of the techniques used for simplifying complex clinical data, as it reduces complexity yet preserves crucial information by identifying the most important aspects of patient datasets. To reduce the redundancy and increase computing efficiency, PCA is used on the hemodynamic indices, lab test findings, and echocardiographic measures. By converting correlated features into independent components, PCA enhances the predictive accuracy of the ECS-DBN model. The dimension reduction is performed by creating an orthogonal basis vector. The basic steps are followed as outlined. Initially, consider the entire clinical information as a matrix for the estimation and assume that the given matrix has a size of N×M, which has to be converted into a M dimensional dataset that is given in equation 2.

$$InputData_{n*m} = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{11} \\ W_{21} & W_{22} & \cdots & W_{2M} \\ \vdots & \vdots & \cdots & \vdots \\ W_{N1} & W_{N2} & \cdots & W_{NM} \end{bmatrix} = [W_1, W_2, .., W_M] \qquad (2)$$

The M dimensional mean vector is calculated using equation 3.

$$W_{mean} = \left(\frac{1}{M}\right) \sum_{j=1}^{M} W_j \qquad (3)$$

The covariance matrix of the clinical dataset is given in equation 4.

$$ConvMat = \begin{bmatrix} Convb_{11} & Convb_{12} & \cdots & Convb_{1M} \\ Convb_{21} & Convb_{22} & \cdots & Convb_{2M} \\ \vdots & \vdots & \cdots & \vdots \\ ConvM_1 & ConvM2 & \cdots & Convb_{MM} \end{bmatrix} \qquad (4)$$

Eigenvalues and eigenvectors are calculated and sorted from the matrix. The obtained eigenvector is used to transform the matrix into a new subsequence. In the above manner, the dimensionality is optimized in the GHW dataset.

Efficient Cockroach Swarm-tuned Deep Belief Network (ECS-DBN) model for readmission prediction

To enhance the prediction of HF readmissions, the research implements the advanced model named DBN with ECSO approach. The potential of DBN to detect intricate correlations between clinical variables allows for earlier treatments to lower readmission rates, while the prediction accuracy and generalization capacity of the model is directly impacted by ECS.

Deep Belief Network (DBN)

To improve the accuracy of 30 and 60 days readmission prediction for patients with HF a suitable method called DBN is utilized. The stacks of Restricted Boltzmann Machines (RBMs) comprise the DBN kind of unsupervised DL model, which successfully develops structured representations of features from clinical data. Every layer of RBM that together build up the DBN model conveys substantial trends from high-dimensional clinical information. There are different functions involved in the DBN. Initially, the DBN is pre-trained using unsupervised training, and the fundamental characteristics in the data are captured by each RBM, as it attempts to recreate the information it provides. After the initial process, the RBM uses the output of the hidden layer of the previous RBM as its input. Multiple RBMs are stacked in effect by repeating that approach with each level. Following the stacking and pre-training of each RBM, supervised learning is used to fine-tune the entirety of the network using labeled information. For improving the network's capacity for certain tasks, such as prediction or classification and adjusting network configurations using methods like backpropagation (BP) fine-tuning is essential. To reduce prediction errors, the BP adjusts the weighting of the entire network depending on labeled information. The combined configurations of RBM energy distribution could be calculated from equation 5.

$$F(u,g,\theta) = -\sum_{ji} X_{ji} u_{ji}\, g_j - \sum a_j\, u_j - \sum b_i g_i \qquad (5)$$

The ECS-DBN parameters are determined by θ=(X,a,b). The bias of visible and hidden components is represented by b_i and a_j respectively. The Boltzmann distribution determines the aggregate distribution of probability, which is given in equation 6.

$$O_\theta(u,g) = \frac{1}{Y(\theta)} \exp\left(-F(u,g,\theta)\right) \qquad (6)$$

Where Y(θ) is the normalization factor. The binary form of the hidden component with the visible component that is i set to 1 with the probability is given in equation 7.

$$O(g_i = 1|u) = sigmoid\left(\sum_j X_{ji} u_j + b_i\right) \qquad (7)$$

The binary form of the visible component with the hidden component that is i set to 1 with the probability is calculated by equation 8.

$$O\left(u_j = 1\middle|g\right) = sigmoid\left(\sum_i X_{ji} g_j + a_i\right) \qquad (8)$$

By identifying significant trends in complicated clinical data, the predictive accuracy is increased by DBN's layered structure, which efficiently detects undetected associations and improves adjusting hyper parameters which guarantees lower computing costs and quicker convergence.

Efficient Cockroach Swarm optimization (ECSO)

The ESCO is the global optimization technique that is inspired by the behavior of the cockroaches. To organize and make decisions, the cockroach colonies can communicate through the process called pheromones. Equation 9 is offered on how location status is conveyed using swarming activities.

$$Z_i = \begin{cases} Z_i + \tau \times \text{rand} \times (M_i - Z_i)\ \forall\ M_i \neq Z_i \\ Z_i + \tau \times \text{rand} \times \left(M_j - Z_i\right) \forall\ M_j = Z_i \end{cases} \qquad (9)$$

The personal optimal location is determined by M_i and the global optimal location is calculated by M_j. When the sight is a constant length of perception, the variations in cockroach behavior are measured by equation 10. Additionally, equation 11 offers the enhanced cockroaches' hunger behavior of ECSO. The weighting factors are mentioned by ω.

$$Z_i = \begin{cases} \omega \times Z_i + \tau \times \text{rand} \times (M_i - Z_i)\ \forall\ M_i \neq Z_i \\ \omega \times Z_i + \tau \times \text{rand} \times \left(M_j - Z_i\right) \forall\ M_j = Z_i \end{cases} \qquad (10)$$

$$Z_i = \begin{cases} Z_i + (Z_i - C_i) + \forall\ y_{food}\ if\ rand < y_{hunger} \\ \qquad y_i + rand\ otherwise \end{cases} \qquad (11)$$

https://doi.org/10.56294/mw2023126

Where, the cockroaches changing locations are denoted by (Z_i- C_i) from the Z_i position. y_{food} represents the precise location, y_{hunger} is the hunger threshold calculation, which is a random value, and c_i is a constant that controls the movement's pace at that particular time. Algorithm 1 provides the entire process flow of the ECS-DBN model.

Algorithm 1: ECS-DBN model

Step 1: Data Collection
Load Dataset (GHW)
Step 2: Data Preprocessing
Function Data Preprocessing(dataset):
Data Cleaning
For each record in Dataset:
Remove duplicate and irrelevant data
Standardize categorical variables (comorbidities, medication history)
Normalize continuous variables using Min-Max normalization:
y_{new} = (y - min(y)) / (max(y) - min(y))
Feature Extraction using PCA
Compute M-dimensional mean vector:
W_{mean} = (1/M) * Sum(W_j)
Compute Covariance Matrix:
ConvMat = Cov(W)
Compute Eigenvalues and Eigenvectors
Sort Eigenvalues and transform matrix using selected Eigenvectors
Reduce dataset dimensionality
Step 4: Train the ECS-DBN Model
Function Train_ECS_DBN (train data, labels):
Initialize DBN with RBM
Pre-train RBMs in an unsupervised manner
For each RBM in DBN:
Train RBM using input data
Use hidden layer outputs as input for the next RBM
Fine-tune DBN using supervised learning with backpropagation
Train DBN on labeled data using cross-entropy loss and backpropagation
Return trained DBN model
Step 5: ECSO Optimization
Function ECSO_Optimization (DBN_model):
Initialize the population of cockroaches with random positions
For each iteration:
Update positions based on swarm and hunger behavior:
If local best q_i ≠ current position y_i:
Update $y_i = y_i + \tau$ * rand * ($q_i - y_i$)
Else
Update $y_i = y_i + \tau$ * rand * ($q_j - y_i$)
If rand < y_{hunger}
Move towards data source y_{food}
Optimize DBN hyperparameters using ECS
Return optimized DBN mode
Step 6: Prediction
Function Predict(model, test_data):
Return model(test_data)
Step 7: Evaluation
Compute F1-score, Precision, Recall, and Accuracy for model performance
Main Execution
Dataset = Load GHW dataset
Preprocessed_Data = Data_Preprocessing(Dataset)
Reduced_Data = Feature_Extraction(Preprocessed_Data)
Trained_Model = Train_ECS_DBN(Reduced_Data, Labels)
Optimized_Model = ECS_Optimization(Trained_Model)
Predictions = Predict(Optimized_Model, Test_Data)
Evaluate_Model (Predictions, Ground_Truth)

The novel ESC-DBN model enhances prediction accuracy by combining the strengths of feature extraction, reduces computational complexity, and optimizes model convergence. By incorporating ECS with DBN, the approach ensures robust learning, effective parameter tuning, and better generalization, leading to improved 30 and 60 days HF readmission prediction.

RESULTS

The ECS-DBN model was legalized in the open dataset, and the suitable system configuration and consistency were verified using different performance metrics compared to the traditional approaches that are elaborated below.

System configuration: A high-performance system arrangement is essential for the ECS-DBN model. The software environment consists of Windows 10 or 11, an Intel Core i7 (8th Gen) processor, and 16 GB of RAM with Python 3.9 as the primary programming language. The DL frameworks including TensorFlow, NumPy, and Scikit-learn are all necessary for the suggested model. The optimal performance for enhanced detection activities is ensured by these configurations. The key performance metrics are described in below.

Accuracy One significant indicator for assessing how well the ECS-DBN model that predicts the readmitted and non-readmitted situations properly for HF is achieved by accuracy. A higher accuracy is needed for the effective prediction. Equation 12 is used to calculate the accuracy.

$$Accuracy = \frac{TP+TN}{TN+FN+TP+FP} \quad (12)$$

Precision: the accuracy of the model's positive predictions is measured by the precision that indicates the proportion of correctly identified readmissions among all predicted readmissions. Reliable identification of high-risk patients is ensured by a greater precision value, which represents fewer errors in diagnosis. Equation 13 is utilized to determine the precision.

$$Precision = \frac{TN}{TN+FP} \quad (13)$$

Recall (sensitivity): the model's recall assesses its capacity to accurately detect every instance of readmission in patients with HF. Out of all real positive outcomes, it calculates the percentage of genuine positive results that were correctly predicted. The prompt care is shown by a greater recall value. Equation 14 is employed to measure recall.

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

F1-score: the F1-score is a key indicator that evaluates a model's predictive capability by balancing precision and recall. An excellent model with an ideal trade-off between detecting actual readmissions and preventing incorrect categorization is indicated by a higher F1-score. Equation 15 delivers the calculation for the F1-score.

$$F1-score = 2 \times \frac{recall \times precision}{recall+precision} \quad (15)$$

The ECS-DBN model's effectiveness with the mentioned metrics is evaluated and compared with the existing method called Cost Sensitive DNN (CSDNN). Table 1 shows the 30-day readmission for the ESC-DBN with the cost sensitive DNN and the graphical representation is offered in figures 2 (a) and (b).

Table 1. ESC-DBN outcomes for 30-days and 60-days readmission

Readmission (days)	Methods	Accuracy	Precision	Recall	F1-Score
30	CSDNN	0,89	0,89	0,26	0,44
	ESC-DNN (Proposed)	0,94	0,95	0,52	0,61
60	CSDNN	0,92	0,93	0,50	0,54
	ESC-DNN (Proposed)	0,96	0,97	0,63	0,65

Although the CSDNN achieved higher accuracy and precision, it had difficulties in feature learning and reliability concerns. The recommended ESC-DNN model achieved 0,94 and 0,96 (accuracy) demonstrating its superiority and effective feature learning in readmission prediction. By achieving 0,95 and 0,97 (precision) ESC-DBN effectively reduces inaccurate results, increasing the trustworthiness of estimations. The value 0,52 and 0,63 (recall) greatly outperformed the existing DNN that reduces neglected scenarios by capturing more real readmissions compared to the cost sensitive, DNN, and 0,61 and 0,65 (F1-Score) obtained the increased class performance.

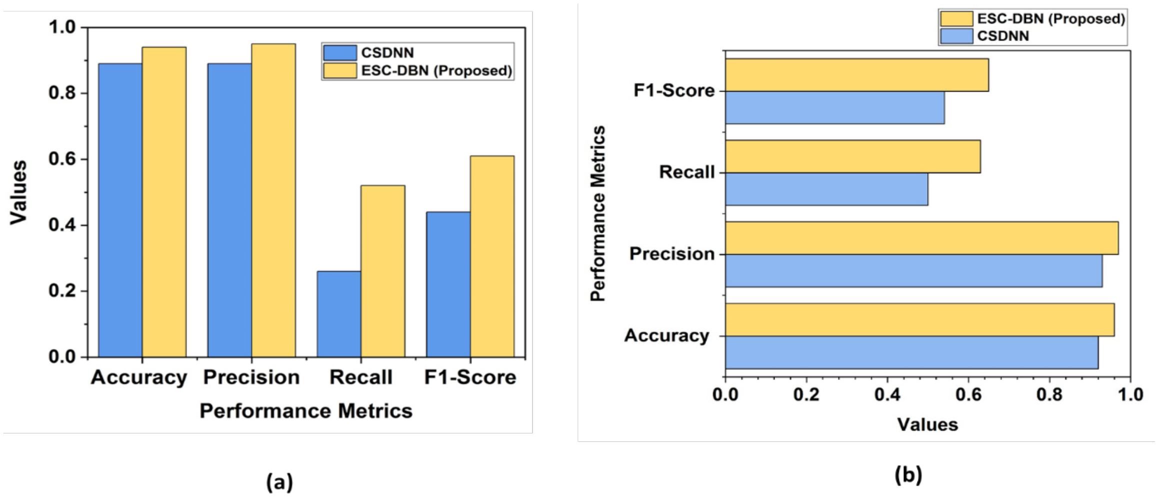

Figure 2. Esc-DBN findings of (a) 30 days and (b) 60 days readmission

DISCUSSION

The CSDNN demonstrated some limitations in identifying readmitted patients due to its suboptimal performance. In contrast, the ESC-DBN model confirmed significant improvements and offered enhanced predictive capabilities. By effectively minimizing inaccurate probabilities, it strengthened the prediction reliability and reduced errors. Additionally, the model significantly improved the recall and ensured the better detection of actual readmissions while lowering missed cases. Its balanced classification approach improved the overall performance and made it a more dependable and efficient tool for readmission prediction. The ESC-DBN model's ability to refine forecasting accuracy and maintain class balance ensures robust decision-making in clinical settings. Its adaptability and efficiency promise a powerful alternative model to conventional models, providing healthcare professionals with a reliable method for identifying high-risk patients.

CONCLUSIONS

The research underscored the effectiveness of ESC-DBN in enhancing the prediction accuracy of HF readmission patients. The ECS-DBN model, optimized using ECS, effectively processes complex clinical datasets of GHW and improves the accuracy in forecasting two categories of 30 and 60-day readmissions. To ensure the superior ESC-DBN model's performance, the technique encompassed the preprocessing approaches like data cleaning and min-max normalization. Additionally, the dimensionality reduction is completed with the help of the PCA. With the layered RBM, the DBN improved the prediction reliability by identifying complicated trends in clinical information. By optimizing DBN hyperparameters, the ECS optimization approach decreased the computing costs and accelerated the convergence. The findings of 30 and 60-day readmission demonstrated the efficient way of the model handling in high-dimensional clinical data compared to conventional models. ECS-DBN highlighted the importance of successful HF treatment for 60 days in terms of accuracy (0,96), precision (0,97), recall (0,63), and F1-Score (0,65). The ECS-DBN model is effective; it mostly involves the comprehensiveness and quality of clinical characteristics. Predictions may be impacted by inadequate or inaccurate information. Further research will be implementing the model into the electronic health record (EHR) environments that could facilitate seamless and automated readmission risk assessment and improve patient management.

BIBLIOGRAPHIC REFERENCES

1. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. European journal of heart failure. 2020 Aug;22(8):1342-56. https://doi.org/10.1002/ejhf.1858

2. Al-Tamimi MA, Gillani SW, Abd Alhakam ME, Sam KG. Factors associated with hospital readmission of heart failure patients. Frontiers in pharmacology. 2021 Oct 11;12:732760. doi: 10.3389/fphar.2021.732760

3. Wideqvist M, Cui X, Magnusson C, Schaufelberger M, Fu M. Hospital readmissions of patients with heart failure from the real world: timing and associated risk factors. ESC heart failure. 2021 Apr;8(2):1388-97. https://doi.org/10.1002/ehf2.13221

4. Akkineni SS, Mohammed O, Pathiraj JP, Devasia T, Chandrababu R, Kunhikatta V. Readmissions and clinical outcomes in heart failure patients: A retrospective study. Clinical Epidemiology and Global Health. 2020 Jun 1;8(2):495-500. https://doi.org/10.1016/j.cegh.2019.11.002

5. Jiang W, Siddiqui S, Barnes S, Barouch LA, Korley F, Martinez DA, Toerper M, Cabral S, Hamrock E, Levin S. Readmission risk trajectories for patients with heart failure using a dynamic prediction approach: retrospective study. JMIR medical informatics. 2019 Sep 16;7(4):e14756.. doi: 10.2196/14756

6. Choi DJ, Park JJ, Ali T, Lee S. Artificial intelligence for the diagnosis of heart failure. NPJ digital medicine. 2020 Apr 8;3(1):54. https://doi.org/10.1038/s41746-020-0261-3

7. Larburu N, Artetxe A, Escolar V, Lozano A, Kerexeta J. Artificial Intelligence to Prevent Mobile Heart Failure Patients Decompensation in Real Time: Monitoring-Based Predictive Model. Mobile Information Systems. 2018;2018(1):1546210. https://doi.org/10.1155/2018/1546210

8. Barbieri S, Kemp J, Perez-Concha O, Kotwal S, Gallagher M, Ritchie A, Jorm L. Benchmarking deep learning architectures for predicting readmission to the ICU and describing patients-at-risk. Scientific reports. 2020 Jan 24;10(1):1111. https://doi.org/10.1038/s41598-020-58053-z

9. Najafi-Vosough R, Faradmal J, Hosseini SK, Moghimbeigi A, Mahjub H. Predicting hospital readmission in heart failure patients in Iran: a comparison of various machine learning methods. Healthcare informatics research. 2021 Oct 31;27(4):307-14.. DOI: https://doi.org/10.4258/hir.2021.27.4.307

10. Awan SE, Bennamoun M, Sohel F, Sanfilippo FM, Dwivedi G. Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. ESC heart failure. 2019 Apr;6(2):428-35. https://doi.org/10.1002/ehf2.12419

11. Sohrabi B, Vanani IR, Gooyavar A, Naderi N. Predicting the readmission of heart failure patients through data analytics. Journal of Information & Knowledge Management. 2019 Mar 21;18(01):1950012. DOI: 10.1142/S0219649219500126

12. Hung M, Lauren E, Hon E, Xu J, Ruiz-Negrón B, Rosales M, Li W, Barton T, O'Brien J, Su W. Using machine learning to predict 30-day hospital readmissions in patients with atrial fibrillation undergoing catheter ablation. Journal of personalized medicine. 2020 Aug 9;10(3):82. doi:10.3390/jpm10030082

13. Hu Z, Du D. A new analytical framework for missing data imputation and classification with uncertainty: Missing data imputation and heart failure readmission prediction. PLoS One. 2020 Sep 21;15(9):e0237724. https://doi.org/10.1371/journal.pone.0237724

14. Wang Z, Zhu Y, Li D, Yin Y, Zhang J. Feature rearrangement-based deep learning system for predicting heart failure mortality. Computer methods and programs in biomedicine. 2020 Jul 1;191:105383. https://doi.org/10.1016/j.cmpb.2020.105383

15. Allam A, Nagy M, Thoma G, Krauthammer M. Neural networks versus Logistic regression for 30 days all-cause readmission prediction. Scientific reports. 2019 Jun 26;9(1):9277. | https://doi.org/10.1038/s41598-019-45685-z

16. https://www.kaggle.com/datasets/programmer3/ghw-heart-failure-readmission-prediction-dataset?select=GHW_HeartFailure_Readmission_Combined.csv

17. Wang H, Cui Z, Chen Y, Avidan M, Abdallah AB, Kronzer A. Predicting hospital readmission via cost-sensitive deep learning. IEEE/ACM transactions on computational biology and bioinformatics. 2018 Apr 16;15(6):1968-78. DOI: 10.1109/TCBB.2018.2827029

FINANCING
No financing.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Data curation: Deepak Kumar Parhi, Pooja Varma, Supriya Awasthi.
Methodology: Deepak Kumar Parhi, Pooja Varma, Supriya Awasthi.
Software: Deepak Kumar Parhi, Pooja Varma, Supriya Awasthi.
Drafting - original draft: Deepak Kumar Parhi, Pooja Varma, Supriya Awasthi.
Writing - proofreading and editing: Deepak Kumar Parhi, Pooja Varma, Supriya Awasthi.