
A Comprehensive Deep Learning Approach for Heart Failure Risk Assessment and Classification

Zainab Hassan¹, Namrata Dhanda¹, Vandana Yadav² and Neha Chauhan¹

¹Department of Computer Science & Engineering, Amity University, Lucknow, Uttar Pradesh, India

²Department of Computer Science & Engineering, Brainware University, Kolkata, West Bengal, India

zainabhassan95561@gmail.com, ndhanda@lko.amity.edu, vandana.yadav771@gmail.com, Neha.chauhan8924@gmail.com

Abstract.

Heart failure is a chronic and progressive cardiovascular condition that contributes substantially to morbidity and mortality worldwide. Early identification of the disease remains difficult because its progression varies among patients and may be associated with a wide range of clinical characteristics. At the same time, the increasing availability of healthcare data has created new opportunities for applying artificial intelligence to medical prediction and classification. In particular, deep learning methods are being investigated for identifying patients who may be at risk of heart failure and for improving the classification of disease-related outcomes.

The present study develops a deep learning-based framework for analysing heart failure using patient clinical records. The approach focuses on identifying patterns within patient data that may be associated with adverse heart failure outcomes. Artificial Neural Networks (ANNs) and other deep learning approaches are considered for their ability to model relationships among multiple clinical variables. The study also discusses commonly used machine learning techniques, including Logistic Regression, Random Forest, Support Vector Machine (SVM), and Gradient Boosting, to provide a broader view of approaches used for heart failure classification.

Model performance can be assessed using measures such as accuracy, precision, recall, F1-score, and AUC-ROC. These metrics provide different perspectives on classification performance and are particularly useful when evaluating medical prediction models. The analysis also emphasizes the importance of data preprocessing and feature preparation before model development, since the quality and consistency of the input data can influence the resulting predictions.

Conventional approaches to heart failure assessment generally involve clinical symptoms, laboratory investigations, and medical imaging. Although these methods remain important in clinical practice, analysing large amounts of patient information can be challenging. Machine learning and deep learning provide computational approaches for examining larger datasets and identifying relationships among multiple variables. Therefore, these techniques may offer useful support for early risk assessment and classification of heart failure, while further validation with larger and more representative clinical datasets remains necessary.

Keywords. Heart disease, deep learning, artificial neural networks, risk assessment, disease classification, hyperparameter tuning.

1. INTRODUCTION

Heart Failure (HF) is a serious and progressive disease characterized by the reduced ability of the heart to pump blood that meets the body's requirements and is related to high morbidity rates. There is an immediate need in the identification of patients with a risk of developing HF or predicting their further course of disease.

At the moment, the task of diagnosing HF includes both clinical examinations, various laboratory tests, and imaging studies. Although these conventional diagnostic methods are effective, they are often resource-intensive and lengthy procedures. More importantly, there are a number of patterns in patients' data that cannot be detected using traditional methods.

At the same time, in combination with the growth of the number of patients with chronic diseases, the introduction of electronic health records and the collection of medical databases has enabled the use of computational methods for assisting medical diagnoses.

Machine learning algorithms have demonstrated great potential for analyzing large sets of patient data and detecting patterns useful in clinical practice. However, most conventional machine learning techniques rely mainly on manually engineered features and have limited abilities in learning complex associations between different attributes.

Deep learning models, on the other hand, enable modeling of complex associations and can help detect patterns that conventional methods fail to detect. Deep learning technology has been introduced as a powerful machine learning approach capable of achieving state-of-the-art results in many challenging pattern recognition tasks. In particular, artificial neural networks showed promising results in clinical applications. For example, they can be used to study the relationship between multiple patient characteristics, such as age, blood pressure, cholesterol levels, heart rate, and other indicators to find patterns associated with the presence of HF. Moreover, a variety of additional features can be added to the model to account for other factors that may affect the development of the disease.

One of the main advantages of using deep learning models is the ability to work with high-dimensional data of different types. In particular, a large amount of data from various sources, including structured data from clinical questionnaires, laboratory tests, physiological measurements, and time series of patient vitals, can be used. The complexity of the interactions between different factors is well known and is one of the challenges for developing and testing appropriate mathematical models. At the same time, artificial neural networks are universal approximators capable of learning complex associations between different attributes. This property is especially useful when working with clinical data since most diseases, including HF, are caused by a combination of different factors. The ability to learn complex interactions between different characteristics makes deep learning models attractive for solving clinical problems.

Another advantage of using deep learning approaches in biomedical research is the ability to simultaneously process multiple types of data. Clinical studies include a vast amount of different types of information, from structured data from clinical questionnaires to time series of patient vital signs and laboratory test results and various imaging data. Each of these types of information usually requires special processing before use in conventional modeling approaches. At the same time, the use of artificial neural networks makes it possible to simultaneously process different types of information using appropriate network architectures. For example, such an approach can be used to process structured data from clinical questionnaires and laboratory tests, time series of patient vitals, and echocardiographic images. Such an integrated approach allows for creating more efficient models since different types of information can be combined in an optimal way to maximize the information content.

Thus, the current project is aimed at developing a systematic deep learning model capable of estimating and classifying the risks of HF. Of primary importance is the selection of a neural network architecture capable of efficiently analyzing the input data while providing sufficient accuracy in predicting the probability of HF. To do this, it is necessary to prepare a set of input data, preprocess them, configure a simple neural network, train it, and evaluate its performance. Applying deep learning approaches to solve clinical problems will allow expanding the capabilities of intelligent medical information systems, improving the accuracy of medical assessments, reducing the time spent on them, and thereby improving the quality of care for patients with chronic diseases such as HF. In particular, this project will demonstrate how deep learning models can be used to predict the development of HF based on patient data.

2. LITERATURE WORK

In recent years, machine learning and deep learning have attracted the attention of researchers who study the prediction of heart failure because of the data on patients' health conditions. Qadeer et al. (2025) compared several machine learning and deep learning models for the task of heart failure prediction and evaluated their performance using clinical data [1]. Their research shows that computational models can help clinicians with analyzing the collected health data, especially when multiple patient-specific attributes need to be taken into consideration. Norouzi et al. (2024) proposed a deep learning model that tackles the survival prediction problem in patients with heart failure by addressing both the notion of frailty and competing risks [2]. The researchers demonstrated the complexity of making clinical predictions about the survival of heart failure patients. Additionally, Scientific Research Publishing (2023) evaluated a combination of machine learning and deep learning methods for the classification of heart failure [3]. Choi

et al. (2017) developed a deep learning model known as the RETAIN, which utilizes a reverse-time attention mechanism for the purpose of creating predictions based on the patient's healthcare data [4]. One of the main advantages of this model is the fact that it allows for distinguishing the most relevant healthcare records, thus enabling the interpretation of the final results. **Rajkomar et al. (2018)** were the first who managed to explore the potential of using deep learning and electronic health records for making various predictions regarding patients' conditions [5]. **Yao et al. (2019)** employed a knowledge-guided convolutional neural network to solve the task of clinical text classification, which helped them address the particular challenge of making predictions based on textual health data [6]. The study illustrates that the analysis of clinical notes can be beneficial because the information that is contained in them is often essential and cannot be found elsewhere. **Krittanawong et al. (2020)** highlight the importance of using machine learning technologies in cardiovascular medicine, as well as discuss their potential benefit for making survival predictions and providing the best patient outcomes [7].

Alaa and van der Schaar (2018) proposed a deep learning-based approach called DeepHit that can be used to predict the survival time of patients with competing risks [8]. **Esteva et al. (2019)** discussed the potential of deep learning algorithms in different areas of medical practice, as well as summarized the ways in which various deep learning-based solutions were used for different prediction and diagnostic tasks [9]. Overall, the authors showed that deep learning technologies can be used in different medical fields despite the variability of input information, task goals, and the final objectives. The use of sequential information about the patient's health condition has attracted the attention of several researchers. For instance, **Lipton et al. (2016)** evaluated recurrent neural networks' ability to learn sequentially structured clinical records and addressed the problem of missing data [10]. **Johnson et al. (2018)** highlighted the role of artificial intelligence and machine learning algorithms in modern cardiology and discussed their potential for clinical use [11]. **Attia et al. (2019)** addressed the possibility of using an artificial intelligence algorithm to analyze electrocardiograms and detect patients who are at risk of developing heart failure [12]. **Cho et al. (2014)** designed an encoder-decoder model that is based on recurrent neural networks and can be used for various sequential tasks [13]. **Miotto et al. (2016)** introduced the Deep Patient framework, which was created using an unsupervised deep learning algorithm that managed to learn the representations of electronic health records and perform multiple prediction tasks [14]. Finally, **Zheng et al. (2021)** reviewed the current research papers and described the deep learning methods used for predicting heart failure readmission [15].

3. USE OF ALGORITHMS

3.1. Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) represent one of the fundamental approaches to deep learning that could be implemented for medical information analysis. The network comprises multiple processing layers of interconnected neurons, where the layers typically comprise an input, a hidden, and an output layer. Each neuron receives input from other neurons through weighted connections and uses an activation function to produce an output signal. For the prediction of heart failure, ANN could be applied to the structured data, which include demographic data, blood pressure, serum creatinine, ejection fraction, and other laboratory metrics. Using multiple variables, it would be possible to explore associations between different risk factors in a multivariate model. Additionally, the network would allow considering numerous patient characteristics related to the condition, thus providing a sufficient exploration of clinical information.

Using the backpropagation procedure, it would be possible to train the network and optimize the weights' distribution in order to minimize the difference between the true and predicted outcome. In this network, the Rectified Linear Units (ReLU) would be used as an activation function for the hidden layer, and a Sigmoid function could be applied to the output layer in order to predict the class probabilities for the binary classification. Overall, ANNs' implementation in the heart failure prediction task would enable building a relatively simple model that could serve as the basis for the solution. Still, it should be noted that the approach's success depends on the data amount and quality, network architecture, and other parameters that should be optimized during the training process.

3.2. Convolutional Neural Network

The CNNs are created for the analysis of the data that have some relationships either in space or time. Hence, CNNs can be helpful in the medical applications of images and signals, particularly, ECG signals. In general, a CNN contains the convolutional layers where different filters are applied to the input data and the patterns and features are identified. These layers can be followed by the pooling operations that allow reducing the dimensions of the data. The CNNs can help in analysing the ECG signals and echocardiographic data in relation to the heart failure research. The CNNs can recognize the patterns that are contained in the forms of the medical data. This recognition cannot be performed using other approaches. In addition, another great benefit of the CNNs is that they are able to learn some features directly from the input data.

Nevertheless, the CNNs have some drawbacks. The training of the CNNs requires a sufficient number of data and can consume much computing resources. The performance of CNNs is influenced by the quality of the input data, network architecture, and training. Therefore, the proper preprocessing and configuration of the CNNs are essential in relation to the heart failure medical data.

3.3. Feed Forward Deep Neural Network

This network is an extended version of the basic Artificial Neural Network (ANN). It has multiple (usually more than 2) hidden layers allowing the model to extract more complex patterns from the input data. Information in the network moves only in one direction from inputs to outputs without backtracking.

Deep Neural Networks are particularly effective in cases where the data consists of many features, such as in the case of clinical data in heart failure patients.

Such networks can easily learn patterns related to different factors that contribute to the condition and better understand the relationships between various inputs. For instance, a DNN can learn to process the information about hypertension, diabetes, and the patient's age and determine how each of these factors taken together affects the likelihood of developing heart failure. At the same time, the complexity of such networks is their disadvantage as too great complexity in relation to the available data will result in overfitting. Thus, proper tuning of the layers and neurons number, learning rate, and other hyperparameters, and the application of dropout regularization are crucial for training such models.

3.4. Recurrent Neural Network

The described below artificial intelligence structures are primarily applied to temporal data as they allow performing the task better than the standard feed-forward neural networks. The main difference consists in having recurrent connections that enable them to learn and retain the hidden state from previous computations. It becomes possible due to the fact that in case of sequential data, the information from earlier steps is required to compute the following ones.

For example, such neural networks can be utilized in medical studies dealing with patients treatment. Specifically, recurrent networks can process the set of patients visits data with particular tests and symptoms as inputs. Moreover, RNNs can operate on the long-term temporal data, which is essential for instance for cardiovascular diseases progress analysis. Nonetheless, vanilla RNNs suffer from a vanishing gradient problem that prevents learning long-term dependencies. For this reason, more advanced Long Short-Term Memory networks and Gated Recurrent Units were developed.

3.5. Long Short Term Memory

LSTM networks, which make use of memory cells and control gates to manage information flow, enable them to learn long-term dependencies better than other neural networks. Such networks have been created to address the shortcomings of standard recurrent neural networks. When it comes to heart failure prediction, LSTM networks can study the progression of the disease over extended periods of time. In other words, the patterns discovered by these networks could be used to extrapolate changes in blood pressure or heart rate that could lead to a heart failure diagnosis. Apart from that,

compared to standard models, LSTM networks tend to have better performance when it comes to prediction tasks. At the same time, their design is more complicated, which means that the implementation and training process could be more challenging.

4. APPLICATIONS OF DEEP LEARNING IN HEART FAILURE RISK ASSESSMENT

1. **Initial Risk Assessment and Preventive Healthcare-** Deep learning algorithms hold great promise for predicting patients likely to suffer from heart failure before the onset of clinical signs. They analyze a number of clinical parameters collectively including laboratory findings, demographics, comorbidities, vital signs as well as patient history. For this reason, they are able to pick up on subtle relationships between different variables that would normally be overlooked by traditional approaches. The value of early risk prediction lies in the fact that it helps physicians initiate preventive measures at the appropriate time. Such preventive measures could range from modifying drug regimens, adopting a healthier lifestyle and follow-ups among others.
2. **Clinical Decision Assistance Systems-** Incorporating deep learning algorithms in hospital management systems will enable practical implementation of this technology in the sphere of medicine. For instance, such algorithms will be able to evaluate patient data and estimate the score or classify the data, thus giving an insight into the difficulty of the case and possible complications. In emergency cases, it may be helpful to predict the need in assistance. Furthermore, automation of such a system will reduce the burden on clinicians and allow them to dedicate their time to other important issues. It should be noted that such algorithms are meant to support medical staff in their work and not to substitute them.
3. **Customized Treatment Strategy-** Heart failure differs from patient to patient since it is influenced by a number of different aspects, including age, genetics, other diseases, and history of treatment among others. With the help of deep learning algorithms, it would be possible to discover the patterns in patient's medical history from which it would be possible to find out what kind of therapy is more suitable for him or her. At the same time, it would be also possible to choose a therapy for a patient that differs from the therapy chosen for everybody else. In this regard, it can be stated that these technologies actually contribute to the personalization of treatment, which is a core of precision medicine.
4. **Ongoing Surveillance and Remote Healthcare Solutions-** Technological improvements in wearable devices plus remote monitoring systems have made it possible to collect continuous physiological data from patients even outside hospitals. Using this real-time information, deep learning algorithms can help clinicians identify patterns that may show worsening of a patient's condition. Changes in heart rate, blood pressure, along with oxygen saturation can be monitored continuously to detect early warning signs. Besides this, such systems may help doctors respond faster before serious complications develop. Continuous monitoring also gives a clearer picture of patient health over time instead of depending only on occasional hospital visits. This can improve patient care as well as support better management of heart-related conditions.
5. **Prediction of Readmission and Mortality-** Readmission rates along with death rates still remain major problems for patients suffering from heart failure. Deep learning algorithms can help in predicting the chances of negative outcomes by studying patterns found in discharge summaries, medications, laboratory reports, plus records of previous hospital admissions. These models may identify high-risk patients earlier compared to traditional approaches. Besides this, proper prediction allows healthcare organizations to plan targeted care programs for patients who need extra attention. Such support can include closer monitoring, medication management, as well as regular follow-up care. In many cases, this may help improve survival rates while also making healthcare delivery more effective and organized.

5. METHODOLOGY

1. To begin with, the proposed method is based on deep learning algorithms that can evaluate and classify the risks of heart failure. In order to prepare the ground for the study, the data set consisting of the records of 300 patients was used. The data set includes such independent variables as age, ejection fraction, serum creatinine, serum sodium, platelet count, and time of the follow-up that are associated with the main diagnosis of heart failure. In this study, the dependent variable was referred to as "DEATH_EVENT," which implies a binary outcome of either being dead or alive ("1" refers to death and "0" refers to being alive).
2. Exploratory data analysis then took place. I have plotted a histogram of each of the clinical parameter. The correlation heatmap was plotted to analyze the feature interaction and whether there exists a correlation level between any two features concerning the target variable.

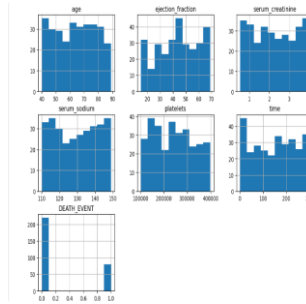


Figure 5.1. Histogram of heart failure features

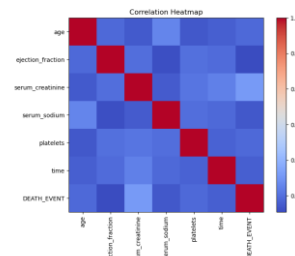


Figure 5.2. Correlation heatmap

3. Once the visualization process was complete, I moved on to the data preprocessing phase. At this stage, I isolated the input variable (X) from the output variable (y). Additionally, an 80/20 split was carried out on the dataset. This implies that 80% of the data was used for training while 20% of the data was used to test the performance of the model. Other than conducting an 80/20 data split, normalization of the data was done by standardizing the range of the variable. This step is important when using a neural network model since it makes the training faster and enhances the model's performance.
4. The classification algorithm was implemented by using Multi Layer Perceptron (MLP) from the Scikit-learn machine learning library using MLP Classifier. The network consists of two hidden layers having 16 neurons in the first layer and 8 neurons in the second one. ReLU was chosen as the activation function for the network whereas the Adam optimizer was used for training the network.

5. Lastly, the model was evaluated using accuracy score along with a confusion matrix. The accuracy score showed the overall prediction performance, while the confusion matrix provided values like true positive, true negative, false positive, plus false negative. These evaluation methods helped in checking the effectiveness of the deep learning classification model step by step.

6. TESTING AND ANALYSIS

The performance of the proposed model was evaluated using an 80/20 split between training plus testing data. Around 80% of the dataset was used for training the model, while the remaining 20% was reserved for testing so that the model could be checked on unseen data samples. Before starting the training process, all variables were standardized to remove differences caused by unequal data scaling. This step helped improve learning efficiency as well as model stability. The proposed neural network architecture was based on a Multi-layer Perceptron containing two hidden layers. ReLU activation functions were applied in the network, whereas the Adam optimization method was used during training to improve the updating of weights inside the model.

7. RESULT AND DISCUSSION

The designed model for heart failure risk classification was analyzed in terms of a well-structured train-test scheme. Training the Multi-Layer Perceptron on the chosen clinical characteristics, the performance metrics on the test sample were estimated. In other words, the level of precision proves that the ability to classify most of the cases properly was demonstrated by the model, implying the existence of the relevant pattern in the training set.

Nevertheless, in order to figure out how the model works in detail, it is crucial to analyze the confusion matrix. One should point out that regardless of several mistakes, the model is able to quite efficiently split patients into two categories depending on whether they have the risk of heart failure or not. Nevertheless, the mentioned mistakes are quite understandable due to the imbalanced data and limited dataset. At the same time, it is worth noting that the results can be supplemented with graphical analysis using histograms and the correlation heatmap. This analysis allows better visualization of the problem since histograms represent the distribution of individual characteristics, while the heatmap reveals the dependencies between variables.

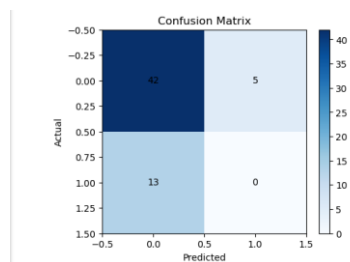


Figure 7.1. Confusion matrix of the classification model

8. CONCLUSION AND FUTURE SCOPE

In conclusion, this study analyzed a deep learning-based algorithm for the evaluation of risk of heart failure by choosing appropriate patient attributes for classification. The artificial neural network has been developed and tested based on the knowledge of the correlation between the patient parameters and his/her health status. It was determined that the proposed algorithm has proven itself quite effective in terms of providing reliable predictions, suggesting its ability to detect important dependencies from the data.

Preprocessing has been used in order to ensure consistency of all patient features in the dataset. Feature scaling, which has been applied to data prior to analysis, allowed for stabilizing the model training process. An additional analysis of the model performance was conducted via metrics such as the calculation of accuracy and confusion matrices, revealing certain inconsistencies, including some misclassifications of patients' status.

The next step to take may include the consideration of other, more complex model architectures and the incorporation of more patient data for predicting better. One may also consider the possibility of combining different medical data sources and building predictive systems that work in real-time.

9. REFERENCES

- [1] Qadeer, M., Ayaz, R., Thohir, M. I., 'Heart Failure Prediction Through a Comparative Study of Machine Learning and Deep Learning Models', Engineering Proceedings, vol. 107, no. 1, MDPI, 2025.
- [2] Norouzi, S., Khormaei, H., Jafarabadi, M. A., Hajizadeh, E., Naderi, N., 'Development of a Deep Learning Model for Survival Prediction in Heart Failure Using Frailty and Competing Risks', Scientific Reports, Nature, 2024.
- [3] Scientific Research Publishing, 'A Novel Approach to Heart Failure Prediction and Classification Through Machine Learning and Deep Learning', 2023.
- [4] Choi, E., Bahadori, M. T., Sun, J., Kulas, J., Schuetz, A., Stewart, W. F., 'RETAIN: Interpretable Predictive Model for Healthcare Using Reverse Time Attention Mechanism', Advances in Neural Information Processing Systems, 2017.
- [5] Rajkomar, A., Oren, E., Chen, K., et al., 'Scalable and Accurate Deep Learning with Electronic Health Records', npj Digital Medicine, 2018.
- [6] Yao, L., Mao, C., Luo, Y., 'Clinical Text Classification with Rule-Based Features and Knowledge-Guided Convolutional Neural Networks', Journal of Biomedical Informatics, 2019.
- [7] Krittawong, C., Virk, H. U. H., Bangalore, S., et al., 'Machine Learning Prediction in Cardiovascular Medicine: Implications for Heart Failure', Journal of the American College of Cardiology, 2020.
- [8] Alaa, A. M., van der Schaar, M., 'DeepHit: A Deep Learning Approach to Survival Analysis with Competing Risks', AAAI Conference on Artificial Intelligence, 2018.
- [9] Esteva, A., Robicquet, A., Ramsundar, B., et al., 'A Guide to Deep Learning in Healthcare', Nature Medicine, 2019.
- [10] Lipton, Z. C., Kale, D. C., Wetzel, R., 'Directly Modeling Missing Data in Sequences with RNNs: Application to Predicting Mortality in Intensive Care Units', Machine Learning for Healthcare Conference, 2016.
- [11] Johnson, K. W., Torres Soto, J., Glicksberg, B. S., et al., 'Artificial Intelligence in Cardiology', Journal of the American College of Cardiology, 2018.
- [12] Attia, Z. I., Noseworthy, P. A., Lopez-Jimenez, F., et al., 'An Artificial Intelligence-Enabled ECG Algorithm for the Identification of Patients with Heart Failure', Nature Medicine, 2019.
- [13] Cho, K., van Merriënboer, B., Gulcehre, C., et al., 'Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation', arXiv preprint, 2014.
- [14] Miotto, R., Li, L., Kidd, B. A., Dudley, J. T., 'Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records', Scientific Reports, 2016.
- [15] Zheng, Y., Peng, H., Zhang, Y., 'Deep Learning Approaches for Predicting Heart Failure Readmission: A Systematic Review', Frontiers in Cardiovascular Medicine, 2021.