Uluslararası Avrasya Sosyal Bilimler Dergisi / *The International Journal of Eurasia Social Sciences*
Cilt/Volume: 16, Sayı/No: 61, Yıl/Year: 2025, ss. / pp.: 1536-1555
E-ISSN: 2146-1961
URL: https://rrpubs.com/index.php/ijoess

Digital transformation in health: An artificial neural network-based prediction model

Muhammed Akif YENİKAYA[1]

[1]Kafkas University, Faculty of Economics and Administrative Sciences, Kars, Turkiye

Araştırma Makalesi/*Research Article*		**DOI:** 10.70736/ijoess.635
Gönderi Tarihi/ *Received:* 11.04.2025	Kabul Tarih/ *Accepted:* 09.09.2025	Online Yayın Tarihi/ *Published*: 15.09.2025

Abstract

Heart diseases remain one of the leading causes of death worldwide, highlighting the urgent need for fast, reliable, and effective diagnostic systems that enable early detection and intervention. This necessity has become more prominent with the increasing integration of digital transformation into the healthcare sector. In this context, advanced technologies such as Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL) have become integral components of clinical decision support systems. This study presents the development of a DL model based on Artificial Neural Networks (ANN) for the early diagnosis of heart disease. The dataset, comprising clinical and demographic characteristics of 303 individuals, was obtained from the open-source UCI platform. The data were preprocessed and divided into training and testing sets, and early stopping was applied during training to prevent overfitting. The developed model achieved an accuracy rate of 90.16%, demonstrating superior performance compared to traditional models. These findings suggest that AI and DL-based systems hold significant potential as reliable and effective decision support mechanisms in the healthcare field, particularly in the diagnosis of heart diseases, within the broader framework of digital transformation.

Keywords: Artificial neural network, deep learning, digital transformation, health sector

Sağlıkta dijital dönüşüm: Yapay sinir ağı tabanlı bir tahmin modeli

Öz

Kalp hastalıkları, dünya genelinde en yaygın ölüm nedenlerinden biri olarak erken teşhis ve müdahale ihtiyacını ön plana çıkarmaktadır. Bu durum, özellikle klinik karar süreçlerinde kullanılmak üzere geliştirilecek hızlı, güvenilir ve etkili tanı sistemlerinin önemini artırmaktadır. Son yıllarda yaşanan dijital dönüşüm, sağlık sektörünü derinden etkilemiş ve bu süreçte yapay zekâ (YZ), makine öğrenmesi (MÖ) ve derin öğrenme (DÖ) gibi ileri teknolojiler klinik karar destek sistemlerinin ayrılmaz bir parçası haline gelmiştir. Bu çalışma kapsamında, kalp hastalıklarının erken tanısına yönelik olarak yapay sinir ağı (YSA) tabanlı bir derin öğrenme modeli geliştirilmiştir. UCI açık veri platformundan elde edilen 303 bireye ait klinik ve demografik özelliklerden oluşan veri seti üzerinde yürütülen bu çalışmada, veriler ön işleme sürecine tabi tutulmuş ve model eğitiminde erken durdurma yöntemi uygulanmıştır. Geliştirilen modelin doğruluk oranı %90,16 olarak elde edilmiş ve geleneksel modellere kıyasla daha yüksek başarı sergilediği görülmüştür. Elde edilen bulgular, dijital dönüşüm sürecinde YZ ve DÖ tabanlı sistemlerin, özellikle kalp hastalıklarının tanısında, güvenilir ve etkin karar destek mekanizmaları olarak önemli bir potansiyele sahip olduğunu göstermektedir.

Anahtar Kelimeler: *Yapay sinir ağı, derin öğrenme, dijital dönüşüm, sağlık sektörü*

Sorumlu Yazar/ Corresponded Author: Muhammed Akif YENİKAYA, **E-posta/ e-mail:** akif.yenikaya@kafkas.edu.tr

INTRODUCTION

Digital transformation is triggering a fundamental transformation in global healthcare systems, reshaping the delivery, management and patient experience of healthcare services. Ranging from electronic health records to artificial intelligence (AI)-assisted diagnostic systems, from telemedicine to mobile health applications, this transformation enables faster, more accessible and personalized delivery of services. Especially in the management of health problems such as heart diseases, which are complex and have high mortality rates, digital health technologies offer significant advantages in critical areas such as early diagnosis, treatment follow-up and patient risk analysis (Mas et al., 2023).

The World Health Organization reports that approximately 17.9 million people die each year due to cardiovascular diseases, with the majority of these deaths (approximately 85%) caused by conditions such as coronary artery disease and stroke (World Health Organization, 2021). Heart diseases are the primary cause of death worldwide, and early and accurate diagnosis is often difficult due to complex clinical symptoms, limited resources and subjective assessments (Bachuk, 2024). At this point, AI and deep learning (DL) technologies make significant contributions to diagnostic processes by providing decision support mechanisms with high accuracy rates in areas such as ECG data analysis, image processing, risk prediction and patient classification (Jahankhani & Kendzierskyj, 2019).

AI and DL are among the most effective components of the digital transformation process in the healthcare sector, significantly transforming traditional diagnosis and treatment processes. The use of AI in healthcare offers great opportunities to improve the accuracy and speed of clinical decision support systems, especially in analyzing large volumes and complex datasets (Topol, 2019). DL provides high accuracy in medical imaging, disease classification and early diagnosis applications thanks to its capacity to learn complex patterns, especially through architectures such as artificial neural network (ANN) (Esteva et al., 2019; Litjens et al., 2017).

DL methods are also widely used in the prediction of heart diseases and have been shown to perform better than traditional logistic regression or decision trees. Moreover, the early detection capability of these algorithms offers strategic advantages for healthcare systems in terms of extending patient life expectancy and reducing healthcare costs (Shickel et al., 2018). With the digital transformation process, the integration of AI applications into the healthcare field is accelerating, especially DL-based solutions both increase diagnostic accuracy and support clinical decision-making processes.

The main objective of this study is to examine the impact of the use of AI in the healthcare sector on heart disease prediction in the digital transformation process. In particular, a decision support system for early diagnosis of heart diseases is proposed through a model developed using ANN. In this direction, the ANN developed using advanced techniques such as early stopping, dropout and batch normalization to increase the accuracy of the model will be analyzed and compared with the traditional model.

One of the key contributions of this study lies in its demonstration of a practical application of digital transformation within the healthcare domain. At the same time, it highlights how DL architectures can be fine-tuned through the adjustment of various parameters to enhance model performance. Furthermore, the results offer valuable insights into how AI-based systems can be improved to deliver more reliable and effective support in clinical decision-making processes.

Digital transformation in the healthcare sector

Digital transformation is not only an innovative process based on the use of technological tools, but also a comprehensive change that fundamentally affects the way organizations do business, service production processes and decision-making approaches. The healthcare sector is one of the areas directly affected by this transformation and experiences the benefits of digitalization most intensely (Agarwal et al., 2010). Especially in the post-pandemic period, the trend towards digital solutions has rapidly increased in many areas such as resilience of health systems, decision processes that require speed and accuracy, data analytics, telemedicine and patient follow-up.

At the center of this transformation are technologies such as big data, AI, Internet of Things (IoT), cloud computing and mobile health applications (Reddy et al., 2019; Taylor et al., 2020). Digitalization of medical data supports healthcare professionals to make accurate and timely decisions, while improving the patient's experience and providing cost-effective solutions. Thanks to digital transformation, previously manual processes have been automated; diagnosis, treatment planning and patient monitoring processes have become faster and more reliable (Wang et al., 2018).

In addition, digital transformation requires not only technological evolution but also cultural and managerial change. Increasing the digital literacy of healthcare professionals, managing data security and ethical issues in a healthy way are integral parts of this process (Jung & Berthon, 2009). In particular, the effective integration of AI-based decision support systems allows healthcare systems to become more sustainable and intelligent.

The heart disease prediction system discussed in this study is an example of AI-enabled health informatics solutions, which is a concrete output of digital transformation. The main motivation of the study is to demonstrate how data-based approaches, which have evolved with digitalization, can be used effectively and reliably in clinical decision processes.

Artificial intelligence (AI)

AI is a broad scientific discipline that aims to develop systems that mimic human-like thinking, learning and problem-solving abilities. ML, which falls under this broad field, is based on the ability of systems to learn patterns and make predictions based on historical data. DL, a sub-branch of ML, is based on multilayer ANN and stands out with its potential to process more complex data structures and achieve high accuracy rates.

In recent years, especially in the healthcare sector, the possibilities offered by AI in areas requiring big data analytics and high processing power have become increasingly important. In many application areas such as diagnostic support systems, disease classifications, patient monitoring systems and medical image processing, AI-supported solutions are yielding successful results (Topol, 2019).

Among the most common AI applications in healthcare are ML and DL techniques. These algorithms can perform functions such as predicting the risk of a particular disease, classifying symptoms, or suggesting a diagnosis by learning from patient data (Rajkomar et al., 2019).

Machine learning (ML)

ML, a subfield of AI, is a set of methods that enable machines to learn from data without being explicitly programmed. ML algorithms extract patterns from large amounts of data, allowing them to predict future situations or make decisions. This feature offers great potential, especially in areas such as healthcare, where data is dense, complex and multidimensional (Obermeyer & Emanuel, 2016).

ML in healthcare has many application areas such as disease diagnosis, treatment recommendations, prediction of patient outcomes, medical image analysis and personalized treatment (Beam & Kohane, 2018). Especially in cases where classical statistical methods are insufficient, multivariate and non-linear relationships are in question, ML-based models can produce more successful results. Classification of heart disease, diabetes prediction, cancer diagnosis and drug response models are examples of such applications (Esteva et al., 2017).

The success of ML depends not only on the structure of the algorithm, but also on data quality, preprocessing, and the correct tuning of model parameters. Therefore, different performance measures such as Precision, Recall, F1-Score and Accuracy should be evaluated together. It is also important to apply strategies such as early stopping and cross-validation to avoid problems such as overfitting.

Deep learning (DL)

DL is a subfield of ML that enables ANN to learn more complex relationships from data using multilayer structures. Thanks to this structure, systems can perform hierarchical learning from low-level features to high-level concepts (LeCun et al., 2015). DL is increasingly being used in the field of healthcare, especially thanks to its ability to work with high accuracy rates on large and complex datasets (Litjens et al., 2017).

ANN, which form the basis of DL algorithms, are multilayer mathematical models that are inspired by biological nervous systems and make decisions by processing data. These networks consist of an input layer, one or more hidden layers and an output layer. Each neuron processes the weighted inputs from the previous layer through an activation function and transfers them to the next layer. In this way, the network can perform classification, prediction or regression by learning patterns from training data (Schmidhuber, 2015).

In the field of health informatics, DL has produced groundbreaking results, especially in applications such as medical imaging, analysis of electronic health records, biosignal data processing and disease prediction. For example, convolutional neural networks (CNN) provide high success in tasks such as tumor detection in medical images, while feedback networks offer effective solutions for analyzing time series data (Esteva et al., 2019).

The multilayer fully connected neural network used in this study is one of the most powerful models for classifying diseases with multiple risk factors such as heart disease. The model was supported with techniques such as dropout and batch normalization, thus preventing overfitting and increasing the generalizability of the model. This powerful structure offered by DL makes significant contributions to making clinical decision support systems more accurate, faster and more scalable.

Literature review

In recent years, rapid advances in technology have enabled the integration of informatics-based approaches, especially AI, into the healthcare sector. The use of these technologies offers significant advantages such as reducing the margin of error in diagnostic processes, improving patient outcomes and increasing the overall efficiency of healthcare services. In literature, there are numerous studies in which AI-based systems have been used effectively in the early diagnosis of various diseases and contributed to the strengthening of clinical decision support mechanisms. Especially in high-risk conditions such as cardiovascular diseases, it is emphasized that AI-supported approaches can significantly help physicians in making accurate and timely diagnoses. In this chapter, we aim to present the current state of knowledge in a systematic framework by discussing current studies on the applications of AI, ML and DL techniques in the field of health, especially in the diagnosis of cardiac diseases.

The study by Mohan et al. (2019) focuses on the effective use of hybrid ML algorithms for early diagnosis of heart diseases. In the study, it was revealed that the hybrid model created by combining different ML techniques achieved higher accuracy rates compared to classical methods. In particular, it is reported that this approach, which uses Random Forest and Naive Bayes algorithms together, achieves a high success rate of 88.7% in heart disease prediction.

Jabbar et al. (2013) proposed an alternative model for heart disease prediction by developing an approach based on lazy association classification (LAC). The study demonstrated that the model obtained by integrating association rules and classification techniques is effective in extracting meaningful information from large medical datasets. This approach facilitates the decision-making process and contributes to clinical decision support systems, especially by generating high-quality rule sets.

Choi et al. (2016) examines the use of Recurrent Neural Networks (RNN) in the early detection of heart failure with data obtained from electronic health records (EHR). The study found that RNN models outperformed traditional machine learning methods in predicting the onset of heart failure thanks to their ability to process time series data. These findings suggest that deep learning models provide an advantage in evaluating time relationships in medical prediction models.

Kumar et al. (2018) used eight different algorithms such as Decision Tree, J48, Logistic Model Tree, Random Forest, Naïve Bayes, K-Nearest Neighbor (KNN), Support Vector Machines (SVM) and Nearest Neighbor for prediction of heart diseases. In their study, it is stated that the prediction accuracy increases as the number of attributes (features) used increases. The main objective of their research is to compare the prediction performance of these data mining and machine learning algorithms on heart diseases and to determine the most effective techniques.

Jiang et al. (2017) comprehensively examine the history, current state and future potential of AI in the healthcare sector. It discusses how AI has revolutionized the analysis of health data and its integration into clinical decision support systems. In particular, applications of AI in areas such as cancer, neurology and cardiology are highlighted. Similarly, the study by Esteva et al. (2019) focuses on the applications of deep learning techniques in healthcare. This study examines the impact of deep learning methods such as computer vision, natural language processing and reinforcement learning in areas such as medical imaging, electronic health records and robotic surgery. The researchers discuss how deep learning can be integrated into medical applications and how these techniques can revolutionize healthcare.

Ripan et al. (2021) developed a heart disease prediction model enhanced with a k-means clustering-based anomaly detection method. In experiments conducted on the UCI dataset, classification was performed using KNN, SVM, NB, RF, and LR algorithms after preprocessing with k-means. This approach increased accuracy rates through anomaly removal.

Xie et al. (2020) proposed a fast k-means clustering algorithm in a big data environment, improving performance and accuracy. The system, tested with image data, was found to be more efficient than traditional k-means, and it was emphasized that the acceleration conditions of the algorithm need to be analyzed in more detail.

Muhammad et al. (2020) developed a model based on smart computing using Cleveland and Hungarian data sets with the aim of early and accurate diagnosis of heart disease. The performance of algorithms such as Logistic Regression, Random Forest, XGBoost, SVM, and Decision Trees was compared, and the Extra Trees and Gradient Boosting models in particular yielded successful results.

Ahamad et al. (2023) investigated the effect of optimal hyperparameter tuning on heart disease prediction. Various algorithms such as RF, KNN, GB, NB, LR, and SVM were tested using the UCI, Cleveland, and Kaggle datasets, and model accuracies were improved using ensemble methods. However, it was emphasized that feature selection strategies did not prioritize the most meaningful features of the dataset, drawing attention to shortcomings in real-time system integration.

Rodríguez and Nafea (2024) compared centralized and federated learning algorithms on the UCI dataset, which consists of data from 920 patients at four hospitals in the US, Hungary, and Switzerland. In the study, the highest accuracy rate of 83.3% was achieved using the SVM algorithm in a centralized structure, and the model's explainability was ensured through SHAP analysis. Using the federated SVM method, an accuracy rate of 73.8% was achieved while maintaining privacy. This study serves as an important reference for privacy-preserving and highly explainable screening models for heart disease prediction.

Jan et al. (2018) developed an ensemble model approach by combining five different classification methods, including SVM, artificial neural network, Naive Bayes, regression analysis, and random forest, to predict the likelihood of recurrence of cardiovascular diseases. The study utilized the Cleveland and Hungarian datasets from the UCI repository; specifically, it was determined that the random forest algorithm demonstrated the highest performance with a 98.12% accuracy rate. These results highlight that ensemble models offer a powerful alternative for disease prediction.

García-Ordás et al. (2023) proposed an innovative model that combines deep learning methods with feature augmentation techniques for the early detection of cardiovascular diseases. The study considered numerous variables such as age, gender, cholesterol, blood sugar, and heart rate, and overcame the challenges of this complex, multi-variable structure through deep learning, which is difficult to address with classical methods. The proposed method outperformed other approaches in the existing literature by 4.4% and achieved a high success rate of 90% accuracy. This result demonstrates that artificial intelligence-based approaches can be an effective tool for the early diagnosis of common and life-threatening heart diseases.

Recent studies have shown promising results in predicting heart disease using machine learning and deep learning techniques. However, a significant portion of these studies do not sufficiently incorporate systematic regularization strategies and comprehensive architectural optimization processes. In response to these shortcomings, the current study adopts a more structured and performance-oriented approach by applying advanced regularization techniques such as early stopping and batch normalization, thereby increasing the model's generalizability and reducing the risk of overfitting. Additionally, the use of the Adamax optimization algorithm, known for its stability in high-dimensional datasets, is another factor that enhances the reliability of the training process. While the majority of previous studies (e.g., Ahamad et al., 2023; Ripan et al., 2021) have focused primarily on ensemble models or hybrid machine learning frameworks, this study proceeds with a simpler deep learning architecture and achieves competitive results with fewer computational resources. In this regard, the proposed model offers a balanced structure between architectural simplicity, prediction performance, and clinical applicability, thereby providing a meaningful and original contribution to the existing literature.

MATERIALS AND METHODS

In this study, a DL-based ANN model for heart disease prediction is developed and compared with a traditional ANN model. In this section, the dataset used in the study, the data preprocessing steps, the architecture of the developed neural network model and the statistical metrics used in the evaluation of the model are explained in detail.

Dataset

In this study, the heart disease dataset obtained from the open-source UCI platform was used (UCI Machine Learning Repository, 2025). UCI datasets are a widely accepted and trusted resource in the academic community due to the accuracy and integrity of the data it contains and the fact that it has been tested in numerous studies by different researchers over the years. Although the dataset is open source, this does not diminish the originality and scientific contribution of the study. This is because this study focuses on the structure of the model, optimization processes and learning strategies rather than the dataset itself. Therefore, the UCI dataset is only a foundation, and the success of the developed model reflects the effectiveness of the AI architecture built on this foundation.

The dataset, whose variables and description are given in Table 1, includes various clinical and demographic characteristics of individuals from different age groups and is based on profiles of individuals at risk of cardiovascular disease. The dataset contains a total of 303 observation units and 13 independent variables, with heart disease status as a binary target variable for each individual (0 = no disease, 1 = disease).

Table 1. Attributes of the dataset

Attribute	Description
age	Age of the patient (in years)
sex	Gender of the patient (0 = male, 1 = female)
cp	Type of chest pain: 0: Typical angina 1: Atypical angina 2: Non-anginal pain 3: No symptoms
trestbps	Blood pressure at rest (in mm Hg)
chol	Serum cholesterol level (in mg/dl)
fbs	Is the fasting blood glucose level above 120 mg/dl? (1 = yes, 0 = no)
restecg	Resting electrocardiogram results: 0: Normal 1: ST-T wave abnormality 2: Probable or definite left ventricular hypertrophy
thalach	Maximum heart rate reached during exercise
exang	Exercise-induced angina (1 = yes, 0 = no)
oldpeak	Exercise-induced ST segment depression compared to resting ST segment depression
slope	ST segment slope during exercise: 0: Ascending 1: Flat 2: Descending
ca	Number of large vessels stained by fluoroscopy (0 to 4)
thal	Thalium stress test result: 0: Normal 1: Fixed defect 2: Reversible defect 3: Unspecified

Source: This table is derived from the heart disease dataset available on the UCI (UCI Machine Learning Repository, 2025)

Data pre-processing

Before starting data analysis, the data set was checked for missing or outlier values, and no missing or outlier data was found. To ensure the effective functioning of neural networks, the input variables were standardized using the StandardScaler method. To accurately assess both the learning success and generalizability of the model, the dataset was split into two parts using the widely accepted practice of an 80% training and 20% testing ratio (Vrigazova, 2021). This ratio allows the model to learn sufficient patterns by allocating a large portion to the learning process, while the remaining 20% is used to objectively measure the model's performance on data it has not seen before. This enables a reliable assessment of how the model can generalize across different datasets while avoiding overfitting.

Model architecture

A deep learning-based ANN model was designed and implemented to solve the classification problem. The model's architecture is structured in layers in a specific order. The first hidden layer uses 128 neurons and the ReLU activation function, followed by a Batch Normalization layer and a Dropout layer with a rate of 0.4 to prevent overfitting. The second hidden layer has 64 neurons and a ReLU activation function, while the third hidden layer has 32 neurons and again uses ReLU activation. The output layer uses a single neuron and a sigmoid activation function, suitable for the binary classification task.

During the training process in ANN models, although the model achieves high success on the training data, its success on the validation data decreases, which may lead to overfitting problems. In order to prevent this problem, early stopping technique was applied in this study. Early stopping is an early intervention mechanism that automatically terminates training if the validation loss of the model does not improve for a certain period.

Thanks to this strategy, the training process continued as long as the model continued to learn, but as soon as no improvement in validation performance was observed, the training was stopped to preserve the generalizability of the model. This is critical to prevent the model from becoming too complex and memorizing, especially in medical datasets with a limited number of observations.

In addition, binary_crossentropy, which is a loss function suitable for the binary classification problem of heart disease prediction, is preferred in this study. Binary_crossentropy provides an optimization objective to improve the accuracy of the model by measuring the difference between the model's predicted probability values and the actual class labels. This function is particularly suitable for cases where a sigmoid activation function is used in the output layer and the target variable is 0 or 1 (Vîrgolici & Virgolici, 2023).

Adamax optimization algorithm was preferred to optimize the weights of the model. Adamax is a derivative of the Adam algorithm based on the L∞ norm and has the capacity to produce more stable results, especially on high-dimensional datasets (Kingma & Ba, 2015). Thanks to its adaptive learning rate, this algorithm enables the model to learn in a more balanced way by updating the weights in small steps and consistently.

The combination of the binary_crossentropy loss function and the Adamax optimization algorithm enhances the model's generalizability, particularly in domains such as medical data where high accuracy and minimal error tolerance are critical, thereby supporting a more stable and effective learning process.

FINDINGS

The DL-based model developed in this study was implemented on Google Colaboratory (Colab), a cloud-based development environment. This open-access platform offered by Google enables the effective execution of ML and DL projects developed in Python programming language. All coding processes, data preprocessing, model training and performance evaluations of the study were carried out using Python. The GPU and TPU support offered by Colab increased the efficiency of the training process by providing both fast processing capacity and hardware independence, especially for DL models working on large data sets. In this context, the technical validity of the research was strengthened by integrating all stages of the developed system with up-to-date software tools and open-source solutions.

Performance indicators for the training process of models

In the plots given in Figure 1, the accuracy and loss values obtained during the training process of the traditional ANN model on the data set used are visualized.

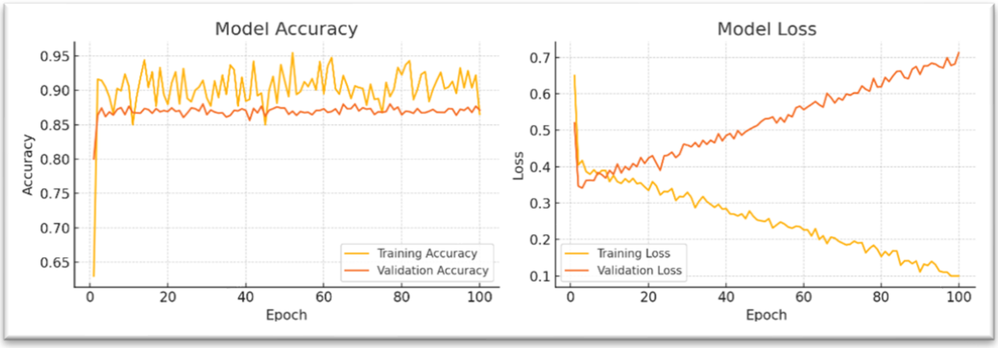

Figure 1. Accuracy-loss plot of the traditional ANN model

While the training accuracy of the model shows an increasing trend, the validation accuracy stabilizes at a certain point and then fluctuates, indicating that the learning capacity of the model is approaching its limit. While the training loss decreases continuously, the validation loss starts to increase after the 30th epoch, indicating signs of overfitting. This indicates that the model fits the training data well, but its generalizability to the validation data is limited.

The confusion matrix presented in Figure 2 shows the details of the model's classification predictions on heart disease diagnosis.

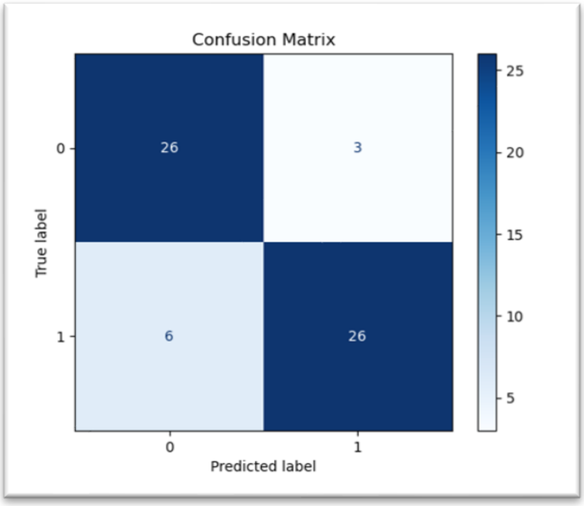

Figure 2. Confusion matrix of the traditional ANN model

When the confusion matrix of the model is examined according to the results obtained:

- True Negative (TN): 26
- True Positive (TP): 26
- False Negative (FN): 6
- False Positive (FP): 3.

The accuracy and loss plots in Figure 3 represent the training process after the early stopping method was applied to the developed neural network model. The accuracy and loss values of the model on the training and validation sets were monitored for 36 epochs.

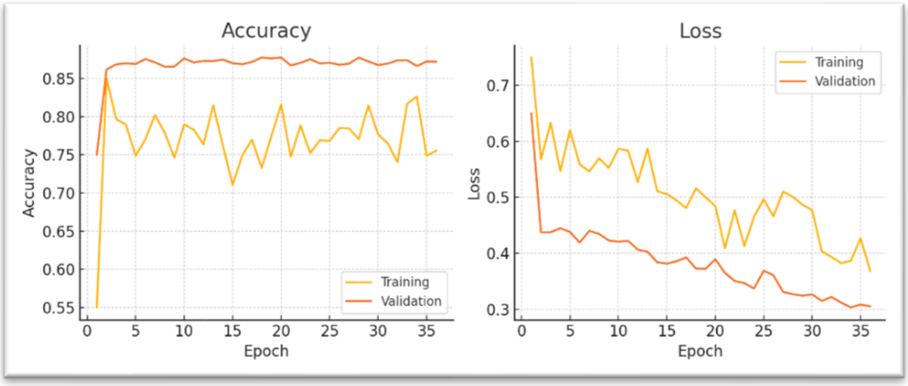

Figure 3. Accuracy-loss plot of the proposed ANN model

The accuracy plot shows a steady increase from the early epochs on the validation set and the model reaches a high level of performance after the 10th epoch. The training accuracy, on the other hand, follows a more fluctuating trend compared to the validation set. This shows that the model gains a generalizable structure without showing an overfitting tendency.

In the loss plot, the validation loss decreased significantly and remained stable throughout the learning process of the model. The training loss, on the other hand, exhibited an upward and downward trend from time to time. This behavior confirms that the early stopping method prevents overlearning and ensures that the training of the model is stopped at the most appropriate point. These results show that the early stopping technique is an effective method to increase the generalizability of the model by preventing it from overlearning.

The confusion matrix presented in Figure 4 represents the performance of the classification process by applying early stopping on the proposed ANN model.

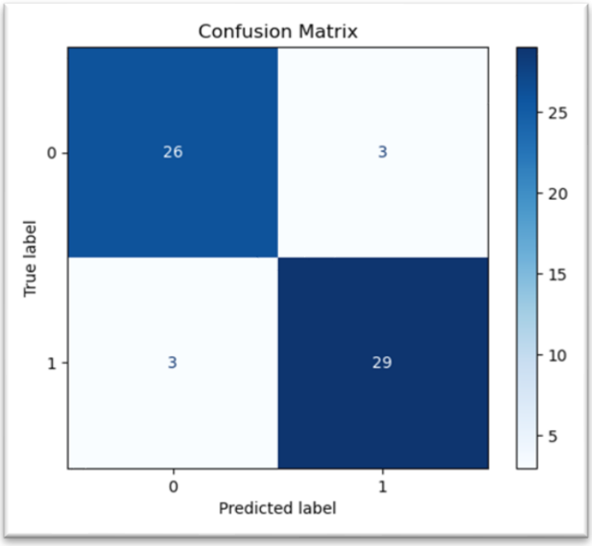

Figure 4. Confusion matrix of the proposed ANN model

When the confusion matrix of the model is examined according to the results obtained:

- True Negative (TN): 26
- False Positive (FP): 3
- True Positive (TP): 29
- False Negative (FN): 3.

Evaluation metrics

One of the most basic tools used to evaluate the performance of the model in classification problems is the confusion matrix. The components of this matrix, True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN), help to understand the accuracy of the model's prediction results.

True Positive (TP) refers to cases where the model correctly predicts a positive state as positive, for example, an individual with heart disease is correctly classified as patient. True Negative (TN) is when the model correctly predicts a negative situation as negative, i.e. a healthy individual is correctly classified as healthy. False Positive (FP) is when a sample that is actually negative is incorrectly classified as positive, i.e. a healthy individual is predicted to be patient. False Negative (FN) means that a sample that should be positive is incorrectly predicted to be negative, i.e. a patient is classified as healthy. These four key elements play a critical role in the calculation of performance metrics such as recall, precision, accuracy and F1-Score. These common metrics used in the literature were used to evaluate the classification performance of the neural network models compared in this study. These metrics are calculated by the following formulas:

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

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

$$F1\text{-}Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

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

Table 2. Calculated performance metrics of the compared models on heart disease prediction

Model	Precision	Recall	F1-score	Accuracy
Random Forest	0.8438	0.8438	0.8438	0.8361
Support Vector Machine	0.9000	0.8438	0.8710	0.8689
K-Nearest Neighbors	0.9333	0.8750	0.9032	0.9015
ANN	0.8965	0.8125	0.8525	0.8525
Proposed ANN	0.9063	0.9063	0.9063	0.9016

This study focuses on comparing the performance differences between the traditional ANN model and the ANN model developed with various regularization techniques such as early stopping. While both models share the same basic architecture, the methods used in the improved model, such as early stopping, dropout, batch normalization, and the Adamax optimizer, have made the learning process more balanced and generalizable. The overfitting problem observed in the traditional model reduces performance on test data, whereas the proposed model delivers significantly better results in terms of accuracy, sensitivity, and F1-score.

Table 2 shows a comparison between the ANN model developed with the early stopping technique and the traditional ANN model, as well as an evaluation of the performance of some machine learning algorithms commonly used in the literature. Such a comprehensive comparison not only reveals the performance of the developed model itself but also increases the validity and reliability of the findings by providing the opportunity to compare it with methods frequently preferred in literature. Identifying the strengths and weaknesses of different algorithms enables valuable conclusions to be drawn about which model may be more suitable under which conditions.

According to the comparison results, although Random Forest (RF) and Support Vector Machines (SVM) models generally performed at an acceptable level, they revealed a weakness that could lead to the misclassification of individuals at risk as healthy, which could be extremely critical from a clinical perspective, particularly due to low recall values. Similarly, although the K-Nearest Neighbor (KNN) algorithm achieved high accuracy rates, it demonstrated limited generalizability in the face of complex data structures. In light of these findings, the ANN model, enhanced with regularization techniques such as early stopping, dropout, and batch normalization, is a meaningful and effective option for clinical decision support systems, particularly due to its high sensitivity level, which enables it to accurately identify individuals at risk.

DISCUSSION AND CONCLUSION

In this study, an artificial intelligence-based deep learning model for the early diagnosis of heart disease was developed, and the effect of the early stopping technique on model performance was examined in detail. Evaluations based on fundamental performance metrics such as accuracy, precision, sensitivity, and F1-score demonstrated that the early stopping strategy reduces the model's tendency toward overfitting, thereby enhancing its generalizability. While the traditional ANN model achieved 85.25% accuracy, the proposed ANN model with early stopping achieved a higher success rate with 90.16% accuracy. Additionally, the precision value increased from 0.8965 to 0.9063, sensitivity from 0.8125 to 0.9063, and the F1-score from 0.8525 to 0.9063. These increases indicate a general improvement in the model's prediction performance.

The performance of the developed model was also evaluated in comparison with other machine learning algorithms applied to the same dataset. Although the KNN algorithm demonstrated strong performance with 90.15% accuracy and an F1-score of 0.9032, it lagged behind the proposed ANN model. The SVM model, despite its high precision (0.9000), was limited in terms of sensitivity (0.8438), which reduced its overall performance, resulting in an F1-score of 0.8710 and an accuracy of 86.89%. The RF algorithm produced relatively lower results across all metrics, achieving 83.61% accuracy and an F1-score of 0.8438. This comparative analysis demonstrates that the ANN model optimized with the early stopping technique is an effective method that provides more balanced and higher accuracy in predicting heart disease.

These results represent a meaningful advancement compared to previous studies. For instance, Jabbar et al. (2013) reported that their LAC model achieved 90% accuracy, outperforming classical methods such as J48 (73.91%) and Naive Bayes (86%). Similarly, Mohan et al. (2019) reached 88.7% accuracy through a hybrid machine learning framework.

Given these benchmarks, the slight yet meaningful improvement achieved by the proposed ANN model demonstrates its potential to refine existing predictive approaches.

The findings suggest that DL architectures provide a more robust representation of learning capacity than traditional ML methods, enabling more accurate predictions on complex healthcare data. Additionally, the integration of the binary_crossentropy loss function and the Adamax optimization algorithm, along with proper data preprocessing and network regularization techniques, contributed positively to model performance.

As a reflection of the broader digital transformation in healthcare, this study illustrates the concrete applicability of AI and DL technologies in clinical decision support. The developed model has demonstrated high performance not only in accuracy but also in classification quality, validating its relevance for practical use. The feature extraction and adaptive learning capabilities of DL models enable fast, scalable, and reliable analysis of critical health conditions. In this context, AI-based systems hold strong promise in helping healthcare professionals make more accurate and timely decisions.

The model presented in this research thus stands as a valuable foundation for the development of intelligent diagnostic tools, combining methodological simplicity with clinically significant performance. Further investigations should explore enhanced tuning strategies, model interpretability, and real-world validation to maximize the impact and trustworthiness of such AI-driven solutions in healthcare environments.

Recommendations

Although the proposed ANN architecture has demonstrated high accuracy and strong generalizability performance, it should be noted that these results were obtained without a comprehensive hyperparameter optimization process. In this regard, it is recommended that systematic methods such as random search or Bayesian optimization be used in future studies to improve the robustness and performance consistency of the model. Additionally, a comparative evaluation of various network architectures using different optimization algorithms may provide deeper insights into determining the most suitable structures for similar datasets.

However, a model must be tested not only based on technical performance metrics but also under different data conditions. In this context, validating the model with broader datasets that include different patient groups will significantly increase the generalizability of future studies. Furthermore, applying the developed model in a real clinical setting and obtaining feedback from healthcare professionals represents a critical step in evaluating the model's practical functionality and validity.

The integration of elements that will increase user confidence is as important as strengthening the applicability of the model. In this context, comparing alternative deep learning architectures and using methods to increase model explainability (e.g., SHAP, LIME) will increase the transparency of the system and reinforce the confidence of end users in the system.

In addition to all these technical and practical considerations, the integration of the model into the healthcare system must also be facilitated. In this regard, converting the model into mobile or web-based applications will facilitate its integration into the digital transformation process of clinical decision support systems. Such applications will contribute to making the practical benefits of digitalization visible, particularly by providing rapid diagnostic support in primary healthcare services.

REFERENCES

Agarwal, R., Gao, G., DesRoches, C., & Jha, A. K. (2010). Research commentary—The digital transformation of healthcare: Current status and the road ahead. *Information Systems Research*, *21*(4), 796-809. https://doi.org/10.1287/isre.1100.0327

Ahamad, G. N., Shafiullah, Fatima, H., Imdadullah, Zakariya, S. M., Abbas, M., ... et al. (2023). Influence of optimal hyperparameters on the performance of machine learning algorithms for predicting heart disease. *Processes*, *11*(3), 734. https://doi.org/10.3390/pr11030734

Bachuk, O. (2024). Digital transformation in the health care system. Odesa National University Herald. *Economy*. https://doi.org/10.32782/2304-0920/3-101-6

Beam, A. L., & Kohane, I. S. (2018). Big data and machine learning in health care. *Jama*, *319*(13), 1317-1318. https://doi.org/10.1001/jama.2017.18391

Choi, E., Schuetz, A., Stewart, W. F., & Sun, J. (2016). Using recurrent neural network models for early detection of heart failure onset. *Journal of the American Medical Informatics Association*, *24*(2), 361–370. https://doi.org/10.1093/jamia/ocw112

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, ... et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118. https://doi.org/10.1038/nature21056

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... et al. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24-29. https://doi.org/10.1038/s41591-018-0316-z

García-Ordás, M. T., Bayón-Gutiérrez, M., Benavides, C., Aveleira-Mata, J., & Benítez-Andrades, J. A. (2023). Heart disease risk prediction using deep learning techniques with feature augmentation. *Multimedia Tools and Applications*, *82*(20), 31759-31773. https://doi.org/10.1007/s11042-023-14817-z

Jabbar, M. A., Deekshatulu, B. L., & Chandra, P. (2013). Heart disease prediction using lazy associative classification. In 2013 *International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s)* (pp. 40-46). https://doi.org/10.1109/iMac4s.2013.6526381

Jahankhani, H., & Kendzierskyj, S. (2019). Digital transformation of healthcare. *Blockchain and Clinical Trial*. https://doi.org/10.1007/978-3-030-11289-9_2.

Jan, M., Awan, A. A., Khalid, M. S., & Nisar, S. (2018). Ensemble approach for developing a smart heart disease prediction system using classification algorithms. *Research Reports in Clinical Cardiology*, 33-45. https://doi.org/10.2147/RRCC.S172035

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... et al. (2017). Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology*, *2*(4). https://doi.org/10.1136/svn-2017-000101

Jung, M. L., & Berthon, P. (2009). Fulfilling the promise: A model for delivering successful online health care. *Journal of Medical Marketing*, *9*(3), 243-254. https://doi.org/10.1057/jmm.2009.26

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*. https://doi.org/10.48550/arXiv.1412.6980

Kumar, M. N., Koushik, K. V. S., & Deepak, K. (2018). Prediction of heart diseases using data mining and machine learning algorithms and tools. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *3*(3), 887-898.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444. https://doi.org/10.1038/nature14539

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60-88. https://doi.org/10.1016/j.media.2017.07.005

Mas, D., Massaro, M., Rippa, P., & Secundo, G. (2023). The challenges of digital transformation in healthcare: An interdisciplinary literature review, framework, and future research agenda. *Technovation*. https://doi.org/10.1016/j.technovation.2023.102716.

Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*, *7*, 81542-81554. https://doi.org/10.1109/ACCESS.2019.2923707

Muhammad, Y., Tahir, M., Hayat, M., & Chong, K. T. (2020). Early and accurate detection and diagnosis of heart disease using intelligent computational model. *Scientific reports*, *10*(1), 19747. https://doi.org/10.1038/s41598-020-76635-9

Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future—Big data, machine learning, and clinical medicine. *New England Journal of Medicine*, *375*(13), 1216-1219. https://doi.org/10.1056/NEJMp1606181

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *New England Journal of Medicine*, *380*(14), 1347-1358. https://doi.org/10.1056/NEJMra1814259

Reddy, S., Fox, J., & Purohit, M. P. (2019). Artificial intelligence-enabled healthcare delivery. *Journal of the Royal Society of Medicine*, *112*(1), 22-28. https://doi.org/10.1177/0141076818815510

Ripan, R. C., Sarker, I. H., Hossain, S. M. M., Anwar, M. M., Nowrozy, R., Hoque, M. M., … et al. (2021). A data-driven heart disease prediction model through K-means clustering-based anomaly detection. *SN Computer Science*, *2*(2), 112. https://doi.org/10.1007/s42979-021-00518-7

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

Shickel, B., Tighe, P. J., Bihorac, A., & Rashidi, P. (2018). Deep EHR: A survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE Journal of Biomedical and Health Informatics*, *22*(5), 1589-1604. https://doi.org/10.1109/JBHI.2017.2767063

Taylor, K., Properzi, F., Bhatti, S., & Ferris, K. (2020). Digital transformation—Shaping the future of European healthcare. *United Kingdom: Deloitte Centre for Health Solutions*.

Topol, E. (2019). *Deep medicine: how artificial intelligence can make healthcare human again*. Hachette UK.

UCI Machine Learning Repository. (2025). *Heart Disease Data Set*. Retrieved March 25, 2025, from https://archive.ics.uci.edu/dataset/45/heart+disease

Vîrgolici, O., & Virgolici, H. (2023). Predicting Prediabetes Using Simple a Multi-Layer Perceptron Neural Network Model. In *ICIMTH* (pp. 168-171). https://doi.org/10.3233/SHTI230453

Vrigazova, B. (2021). The proportion for splitting data into training and test set for the bootstrap in classification problems. *Business Systems Research: International Journal of the Society for Advancing Innovation and Research in Economy*, *12*(1), 228-242. https://doi.org/10.2478/bsrj-2021-0015

Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change*, *126*, 3-13. https://doi.org/10.1016/j.techfore.2015.12.019

World Health Organization. (2021). Cardiovascular diseases (CVDs). World Health Organization. Retrieved March 25, 2025, from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Xie, T., Liu, R., & Wei, Z. (2020). Improvement of the fast clustering algorithm improved by K-means in the Big Data. *Applied Mathematics & Nonlinear Sciences*, *5*(1).

KATKI ORANI ***CONTRIBUTION RATE***	**AÇIKLAMA** ***EXPLANATION***	**KATKIDA BULUNANLAR** ***CONTRIBUTORS***
Fikir ve Kavramsal Örgü *Idea or Notion*	Araştırma hipotezini veya fikrini oluşturmak *Form the research hypothesis or idea*	Muhammed Akif YENİKAYA
Tasarım *Design*	Yöntem ve araştırma desenini tasarlamak *To design the method and research design.*	Muhammed Akif YENİKAYA
Literatür Tarama *Literature Review*	Çalışma için gerekli literatürü taramak *Review the literature required for the study*	Muhammed Akif YENİKAYA
Veri Toplama ve İşleme *Data Collecting and Processing*	Verileri toplamak, düzenlemek ve raporlaştırmak *Collecting, organizing and reporting data*	Muhammed Akif YENİKAYA
Tartışma ve Yorum *Discussion and Commentary*	Elde edilen bulguların değerlendirilmesi *Evaluation of the obtained finding*	Muhammed Akif YENİKAYA

Destek ve Teşekkür Beyanı/ *Statement of Support and Acknowledgment*

Bu çalışmanın yazım sürecinde katkı ve/veya destek alınmamıştır.

No contribution and/or support was received during the writing process of this study.

Çatışma Beyanı/ *Statement of Conflict*

Araştırmacıların araştırma ile ilgili diğer kişi ve kurumlarla herhangi bir kişisel ve finansal çıkar çatışması yoktur.

Researchers do not have any personal or financial conflicts of interest with other people and institutions related to the research.

Etik Kurul Beyanı/ *Statement of Ethics Committee*

Bu çalışma, kamuya açık ve anonimleştirilmiş veriler kullanılarak gerçekleştirildiğinden dolayı etik kurul onayı gerektirmemektedir.

This study does not require ethical committee approval since it was conducted using publicly available and anonymized data.