Contents lists available at *Dergipark*

Journal of Scientific Reports-A

journal homepage: *https://dergipark.org.tr/tr/pub/jsr-a*

E-ISSN: 2687-6167 ***Number 55, December 2023***

RESEARCH ARTICLE

Receive Date: 11.08.2023 *Accepted Date: 06.10.2023*

Early-stage heart failure disease prediction with deep learning approach

Emin Demir[1*], Ferhat Bozkurt[2], Yusuf Ziya Ayık[3]

[1] *Atatürk University, Institute of Social Sciences, Management Information Systems PhD Program, Erzurum, ORCID:0009-0007-9657-2079*
[2] *Atatürk University, Faculty of Engineering, Department of Computer Engineering, Erzurum, ORCID:0000-0003-0088-5825*
[3] *Atatürk University, Vocational School of Technical Sciences, Computer Technologies Program, Erzurum, ORCID:0000-0002-7857-9417*

Abstract

Cardiovascular diseases rank the highest among diseases in terms of mortality rate and cause millions of deaths every year. Heart failure is a type of cardiovascular disease and its early diagnosis is extremely important for its prevention. It may be vitally important to understand to what extent which body values, characteristics and factors (age, gender, blood pressure, sugar, etc.) affect this disease and to predict whether the individual will have a possible heart attack in the future. In this study, firstly, the correlation level of the relevant body values with the disease is extracted and in the second stage, a method that predicts heart attack with DNN (Deep Neural Network) and CNN (Convolutional Neural Network) deep learning models is proposed. In the study, 918 observations obtained from the kaggle site were used. Firstly, missing data, categorical data, non-numerical features were checked. Then, outliers were cleaned and the relationship of the features in the dataset with the disease state was revealed by feature engineering operations on the data. Finally, deep neural network models were built and the model was trained and hyperparameter adjustment was performed with GridSearhCV to achieve the highest success rate. As a result of the study, Accuracy, Precision, Recall and F1-Score values were found as 0.9375, 0.9629, 0.9176, 0.9397 for DNN and 0.9312, 0.9512, 0.9176, 0.9340 for CNN respectively. The AUC value calculated from the ROC curve was found to be equal to 0.96 in both deep learning models.
Keywords: Deep learning, DNN, CNN, Heart attack prediction

1. Introduction

* Corresponding author. Tel.: +90-533-564-4954.
E-mail address: e_demir@odu.edu.tr

Cardiovascular diseases take nearly 18 million lives each year and are shown as the highest cause of death worldwide. The WHO (World Health Organization) states that the number of deaths due to heart disease will increase even more over the next 10 years. Early diagnosis of such a risky disease is vital to prevent the number of deaths due to the disease [1], [2]. The fact that there are many features to be checked in the diagnosis of heart failure (age, gender, type of chest pain, blood pressure, fasting blood sugar, ECG, etc.) makes the work of physicians seriously difficult [3].

For this reason, decision support systems that shorten the test time of this disease with artificial intelligence methods such as fuzzy logic, machine learning and deep learning and that can help physicians in the diagnosis of the disease have become one of the extremely valuable topics that have been intensively studied in recent years. When the studies carried out in the literature for heart failure prediction are analysed, it is seen that different artificial intelligence approaches and different techniques (pre-processing, extraction, clustering and hyperparameter etc.) have achieved success between 63.33 and 93.44. On the other hand, it is understood from the literature review that the majority of the studies for heart failure prediction are performed with machine learning methods, while the number of similar studies with deep learning is quite low.

In this study, in order to fill this gap in the literature and to help doctors in predicting heart failure, two different models were created with DNN and CNN deep learning methods and the performance of the models was compared with metrics accepted in the literature.

The important contributions of this article are:

- In the study, the kaggle heart failure prediction dataset, which is a combination of 11 common features of the universally accepted Cleveland (n=303), Hungarian (n=294), Swiss (n=123), Long Beach VA (n=200) and Stalog-Kalp (n=270) datasets, was used.
- Two different models were created with DNN and CNN deep learning methods and the performance of the models were compared with metrics accepted in the literature.
- In order to obtain the best result with the created deep learning models, hyperparameter adjustment was performed using GridSearhCV.
- At the end of the study, a model that can achieve Accuracy, Precision, Recall and F1-Score values in all of them in the range of 91% to 96% was proposed.

In the following sections of the study;
Part 2 – Related studies in the literature,
Part 3 – Materials and methods,
Part 4 - Research findings,
Part 5 - Discussion and conclusion
parts are included.

2. Related studies in the literature

At the end of the study, a model that can achieve Accuracy, Precision, Recall and F1-Score values in all of them in the range of 91% to 96% was proposed.

Yan et al. developed a Multi-Layer Perceptron (MLP) based decision support system to support the diagnosis of heart diseases. Using a total of 352 medical records collected from 5 heart patients, over 90% success was achieved. [4]. Das et al. obtained 89.01% accuracy using Statistical Analysis System (SAS) 9.1.3 core software, a neural network ensemble at its core, in the study for the diagnosis of heart disease [5]. In another study conducted by Abdullah and Rajalaxmi for heart disease diagnosis, Random Forest and Decision Tree models were used and 63.33% and 50.67% success was achieved respectively [6]. Shao et al. proposed a hybrid model approach using Logistic Regression, Multivariable Adaptive Regression Splines (MARS), Rough Clustering (RS) and Artificial Neural

Network and were achieved 83.93% success with both MARS-LR and RS-LR [7]. Cihan et al. In their study for the detection of Coronary Artery Disease, Naive Bayes, J48 Decision Tree, Multilayer artificial neural network (ANN) and 1R classification methods were compared with 303 case studies and they were obtained the best result with Multilayer ANN with a rate of 83.498% [8]. In their study, Özmen et al. compared the performance of J48, Logistic Regression, Support Vector Machine (SVM), Naive Bayes, Multilayer Perceptron, Random Forest, Adaboost, Bagging decision Trees and Single Layer Perceptron models considering their advantages and disadvantages. As a result of the comparison, SVM achieved the highest success rate with 89.47% [9]. Ozcan et al. aimed at early detection of heart disease with Support Vector Machine and Forward Propagation ANN using 170 case data and achieved 83.33% success with Forward Propagation ANN and 91.67% success with Support Vector Machine algorithm [10]. Göktaş and Yağanoğlu used Random Forest, K-Nearest Neighbor, SVM and C4.5 decision tree algorithms for heart attack prediction and compared their results. As a result of the study, the highest success rate was obtained in the C4.5 Decision Tree algorithm with an accuracy rate of 83.09%. [11]. In their study for the diagnosis of heart disease, Ekrem et al. achieved 86.88% success with Random Forest classification method by selecting features with particle swarm optimization on 303 cases [12]. In this study by Coşar and Deniz, heart disease early detection algorithm was developed with Random Forest, Logistic Regression and k-NN algorithm. The highest success rate of 88% was achieved with Random Forest [13]. Potur and Erginel applied different feature selection methods on 299 case data and compared Support Vector Machines, Logistic Regression, Multilayer Perceptron, J48 decision tree and Naive Bayes algorithms. The highest success rate of 90% was achieved with Multilayer Perceptron (MLP) [14]. In the study conducted by Gündoğdu for early detection of heart disease, 303 cases were examined and 7 different machine learning algorithms including Linear Discriminant Analysis, Decision Trees, K-Nearest Neighbor, Support Vector Machines, Naive Bayes, Random Forest and Gradient Boosting were compared. The best accuracy rate of 90.02% was obtained with the Random Forest machine learning algorithm [15]. Yilmaz and Yağin in 2021, in the study they aimed to compare the prediction results of Radial Basis Neural Network (RBF) and Multilayer Neural Networks (MLP), they reached an accuracy of 91.1% with MLP and 79.7% with RBF [16]. In the study conducted by Vatansever et al., Decision Tree, Logistic Regression, k-NN, Random Forest, Support Vector Machine and Naive Bayes models were used to predict heart disease with various machine learning algorithms by selecting features with Genetic Algorithm approach. As a result of the study, the highest success rate was obtained with 93.44% in the experiments with the genetic algorithm approach [17]. In the study conducted by Salman and Aksoy, hyperparameter selection was performed with the Random Forest algorithm and then feature selection was performed using the binary particle swarm intelligence method. At the end of the study, 79.66% success rate was achieved [18]. In their study, Coşkun and Kuncan examined 918 cases obtained through kaggle and evaluated the performances of Linear Discriminant Analysis, Support Vector Machines, Gaussian Naive Bayes, Random Forest and Decision Tree machine learning algorithms and found that the highest success rate was 90.76% with Random Forest [19]. In their study, Çil & Güneş compared the performance of Support Vector Machines, Artificial Neural Networks, Naive Bayes, Random Forest algorithms and k-NN machine learning algorithms in detecting heart diseases. As a result, the best success rate was obtained with the Logistic Regression machine learning algorithm with 90.77% [20]. Yilmaz ve Yağin in 2022, compared the performance of Logistic Regression, Support Vector Machine machine and Random Forest learning algorithms for the prediction of coronary heart disease, the highest success was obtained in Random Forest algorithm with 92.90% [21]. Keser and Keskin, in their study, Logistic regression, Gradient Boosting and Artificial Neural Network models were used to predict the survival or mortality of patients with 299 cases and the best prediction result was obtained with the Artificial Neural Network model with 86.67% [22].

3. Material and methods

3.1. Data set

In this study, used the *kaggle* heart failure prediction dataset consisting of 918 observations by combining 11 common features of the universally accepted Cleveland (n=303), Hungarian (n=294), Swiss (n=123), Long Beach VA (n=200) and Stalog (Heart) (n=270) datasets. There are 11 independent and one dependent variables in the data set, and the variables and their explanations are presented in Table 1.

Table 1. Variables in the data set and their descriptions.

Feature	Feature Description	Type
Age	Age of the patient (Years)	Independent
Sex	Gender of the patient (Female/Male)	Independent
ChestPainType	Type of chest pain (4 Values)	Independent
RestingBP	Resting state blood pressure (mm Hg)	Independent
Cholesterol	Serum cholesterol (mg/dl)	Independent
FastingBS	Fasting blood glucose (120 mg/dl > 1 if not 0)	Independent
RestingECG	Resting ECG results (0 normal, 1 ST-T abnormality, 2 possible or definite left ventricular hypertrophy according to Estes criteria)	Independent
MaxHR	Maximum heart rate	Independent
ExerciseAngina	Exercise-induced angina (0, 1)	Independent
Oldpeak	Numerical value measured in depression	Independent
ST_Slope	Slope of the Peak Exercise (0 upward, 1 straight, 2 downward)	Independent
HeartDisease	Probability of a heart attack (0 low, 1 high)	Dependent

The basic statistical information of the non-categorical numerical variables in the data is presented in Table 2. According to Table 2, there are 918 data for each variable. The mean of the age variable is 54. The lowest age is 28 and the highest is 77. The Q3 value of the age variable is less than or equal to 60. In other words, 75% of the age variable is less than or equal to 60. The mean of the Cholesterol variable is 198.79 and the lowest Cholesterol is 0 and the highest is 603. The Q3 value of the variable Cholesterol is less than or equal to 267. The mean of the RestingBP variable is 132.39 and the lowest RestingBP is 0 and the highest is 200. The Q3 value of the RestingBP variable is less than or equal to 140.

Table 2. Basic statistical knowledge of numeric columns.

	n	mean	std	min	25%	50%	75%	max
Age	918	53,51089	9,432617	28	47	54	60	77
RestingBP	918	132,3965	18,51415	0	120	130	140	200
Cholesterol	918	198,7996	109,3841	0	173,25	223	267	603
FastingBS	918	0,233115	0,423046	0	0	0	0	1
MaxHR	918	136,8094	25,46033	60	120	138	156	202
Oldpeak	918	0,887364	1,06657	-2,6	0	0,6	1,5	6,2
HeartDisease	918	0,553377	0,497414	0	0	1	1	1

In the data set, 410 patients (44.66%) did not have a heart attack, while 508 patients (55.34%) had a heart attack (Figure 1).

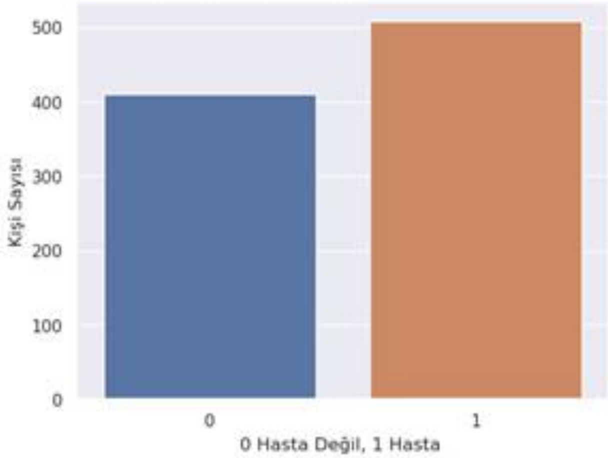

Fig. 1. Distribution of the number of patients in the data set.

3.2. Data preprocessing and feature selection

3.2.1. Data preprocessing and feature selection

In data preprocessing, first, missing data control and categorical and non-numerical features were checked. It was understood that there was no missing data in the data set, but the variables 'Sex', 'RestingECG', 'ExerciseAngina', 'ST_Slope', 'ChestPainType' were understood to be non-numeric categorical data. In order to digitize these variables, the digitization process was carried out using the One-Hot Encoding method since the data set was not very large. As a result of the One-Hot Encoding digitization process, the variables in the dataset are "'Age', 'RestingBP', 'Cholesterol', 'FastingBS', 'MaxHR', 'Oldpeak', 'HeartDisease', 'Sex_F', 'Sex_M', 'RestingECG_LVH', 'RestingECG_Normal', 'RestingECG_ST', 'ExerciseAngina_N', 'ExerciseAngina_Y', 'ST_Slope_Down', 'ST_Slope_Flat', 'ST_Slope_Up', 'ChestPainType_ASY', 'ChestPainType_ATA', 'ChestPainType_NAP', 'ChestPainType_TA'" as increased to 21.

3.2.2. Examination of data distribution and outlier control

In this step, it was examined whether the non-categorical numeric variables ('Age', 'RestingBP', 'Cholesterol', 'FastingBS', 'MaxHR', 'Oldpeak') in the data set were normally distributed. In addition, outlier control was performed for the data.

In this study, Q-Q Plot is used to determine the distribution of the data by visualizing the data. Q-Q Plot performs an assumption check. It is a method based on quantiles to examine whether the data are normally distributed. In the working principle of Q-Q Plot, the data set is sorted and the sorted data and standard normally distributed data are shown on the graph in the coordinate plane. If the two lines overlap linearly at an angle of 45 degrees, it is understood that the data are normally distributed [23], [24].

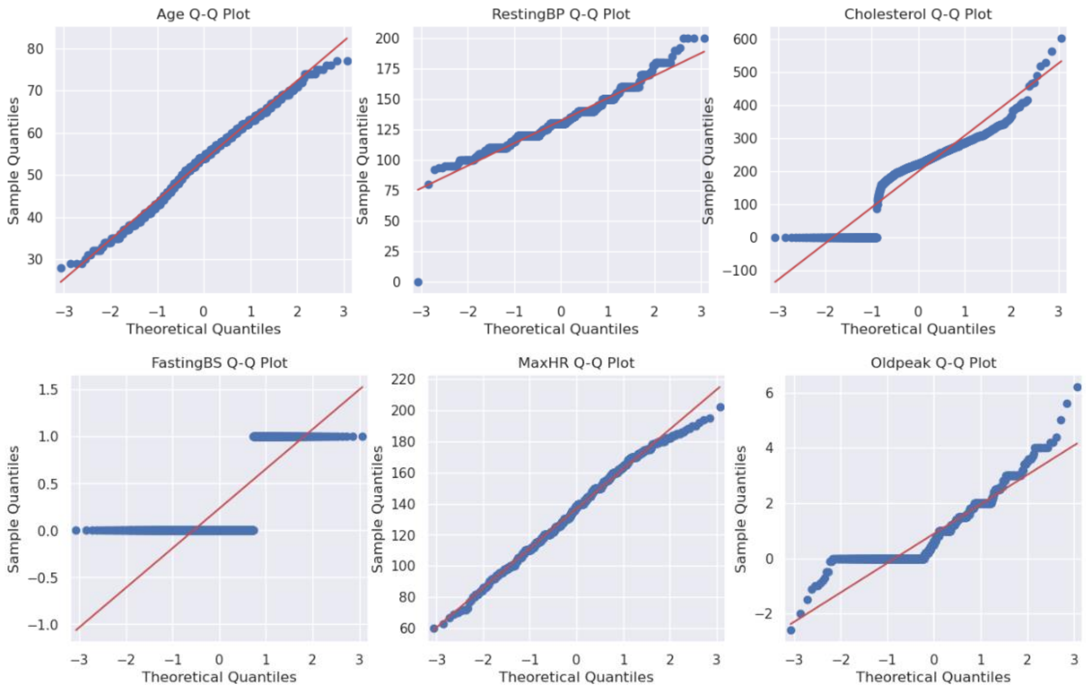

Fig. 2. Q-Q Plot data distribution before excluding outliers for Age, RestingBP, Cholesterol, FastingBS, MaxHR, Oldpeak variables in the dataset.

When the Q-Q Plot presented in Figure 2 is examined, outliers between 0 and 25 are clearly seen in the 'RestingBP' variable. In the variables 'Cholesterol' and 'Oldpeak', the lines parallel to the x-axis mean that the data distribution showed deviations from the normal distribution and this can be interpreted as skewness. As for the 'FastingBS' variable, it is seen that the data are concentrated on two peaks in the graph. However, since the data of the 'FastingBS' variable consists of 1 and 0, there is no skewness in this variable.

There are various methods such as standard deviation, z-score, 1st and 99th percentile to detect outliers. In this study, the standard deviation method was used to detect outliers and exclude them from the data set. In the standard deviation method, the mean of the data set is firstly calculated and then the standard deviation of the data set is found.. A threshold value for outliers is set. When the studies in the literature are examined, it is seen that in the calculation of the threshold value, researchers generally use a multiple of the standard deviation, such as 2 or 3 standard deviations from the mean [25]. In this study, values greater than the final threshold distance were identified as outliers and excluded from the dataset. The Q-Q Plot generated after this process is presented in Figure 3.

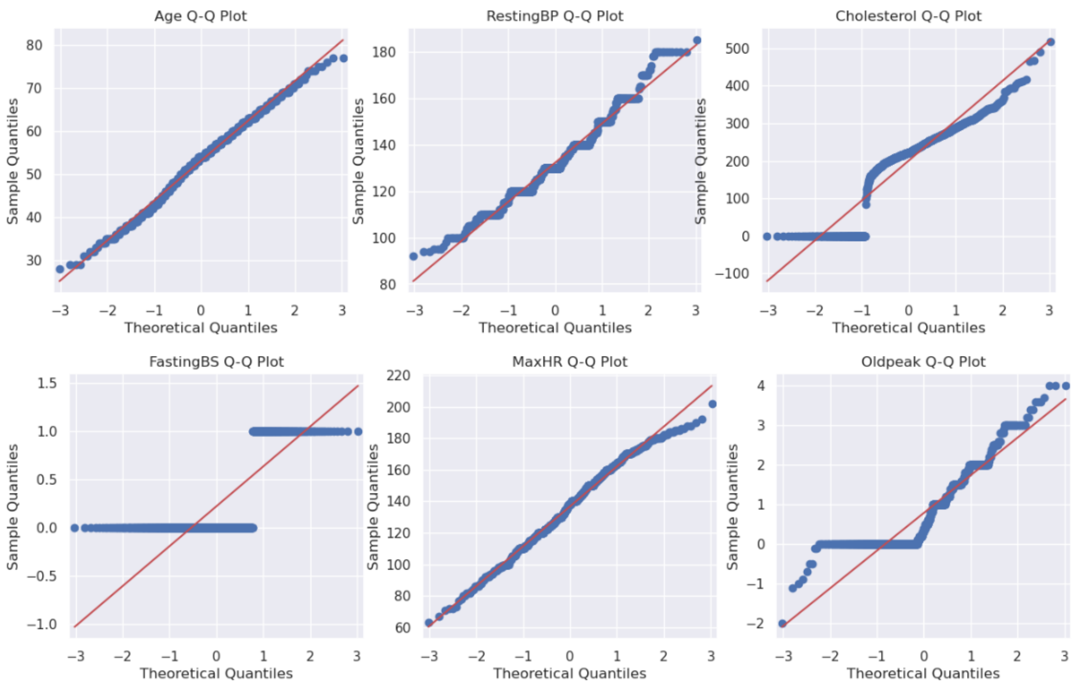

Fig. 3. Q-Q Plot data distribution after excluding outliers in Age, RestingBP, Cholesterol, FastingBS, MaxHR, Oldpeak variables in the dataset.

When the Q-Q graphs presented in Figure 2 and Figure 3 are analyzed comparatively, it is understood that the numerical data are distributed more regularly after the outliers are removed, especially in the 'RestingBP' variable. However, in both figures, 'Cholesterol' and 'Oldpeak' variables still show lines parallel to the x-axis. In order to be understand better whether there is skewness in these variables, the histogram graphs of 'Cholesterol' and 'Oldpeak' variables are given in Figure 4.

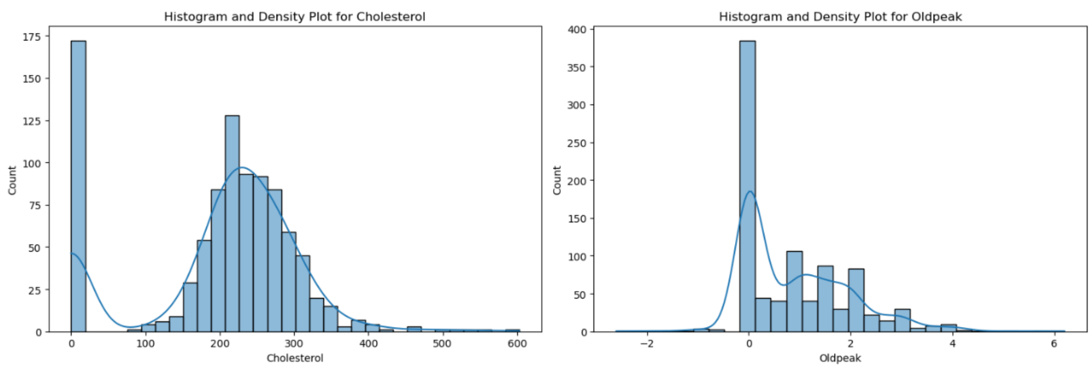

Figure 4. Histogram plot for Cholesterol and Oldpeak variables in the dataset.

Data may be skewed if the data distribution is not symmetrical, if there is skewness on the right side or on the left side, and if the peak value is far from the centre [23]. In line with this explanation, when the histogram graphs presented in Figure 4 are analysed, the peak value increase in the 'Oldpeak' variable can be interpreted as skewness. Similarly, presence of two different peaks in the data set, if the data are concentrated at two points, this means that the data are bimodally distributed. In a bimodal distribution, the data are not concentrated around one value, but around two different values, and this is evaluated as skewness. This is also observed in the variable 'Cholesterol'.

3.2.3. Normalisation and standardisation processes

When the histogram graphs presented in Figure 4 are analysed, it is understood that there is skewness in the data distribution of 'Cholesterol' and 'Oldpeak' variables. After analysing the distribution of the data, the normalisation step was started. Since 'Oldpeak' and 'Cholesterol' variables were found to be skewed, these variables were scaled with MinMaxScaler (Equation 1). StandardScaler (Equation 2) was used for other numeric variables.

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

$$X_{scaled} = \frac{X - \mu}{\sigma} \quad (2)$$

3.2.4. Correlation analyses

In the correlation analysis step, it was observed that the correlation analysis could not be created for the variables 'ChestPainType_TA' and 'ST_Slope_Down'. This may be due to missing data in the variables or fixed data in the variables in question. Since the entire data set was checked at the beginning of the study and there was no missing data, it was checked whether the variables in question took constant value or not, and it was found that these two variables took constant value. For this reason, these two variables were excluded from the correlation analysis since including or excluding these two variables in the analysis would not change the result. Since the representation of 19 variables in the matrix reduces the legibility of the matrix, only numerical variables are included in the matrix representation (Figure 5).

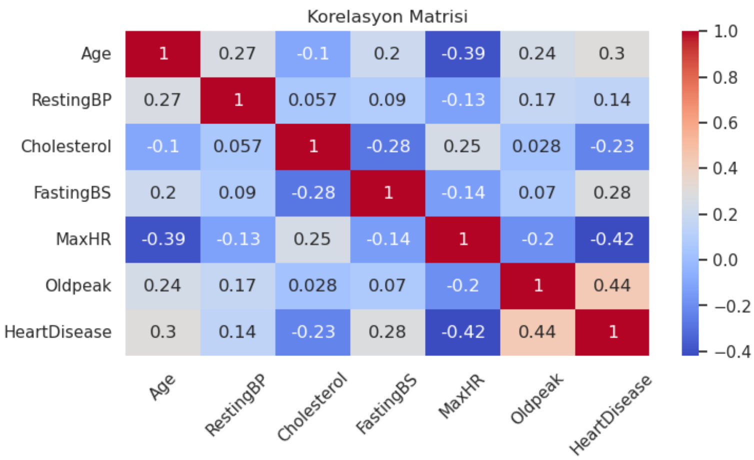

Fig. 5. Correlation matrix of numeric variables in the dataset.

When the correlation matrix of numerical values was analysed, with the variable HeartDisease;

- Oldpeak, RestingBP, FastingBS and Age variables have a positive relationship,
- MaxHR and Cholesterol have a negative relationship,
- The highest relationship level is in Oldpeak and MaxHR variables with 0.44 and -0.42, respectively,
- The lowest relationship level was found in the RestingBP variable with 0.14.

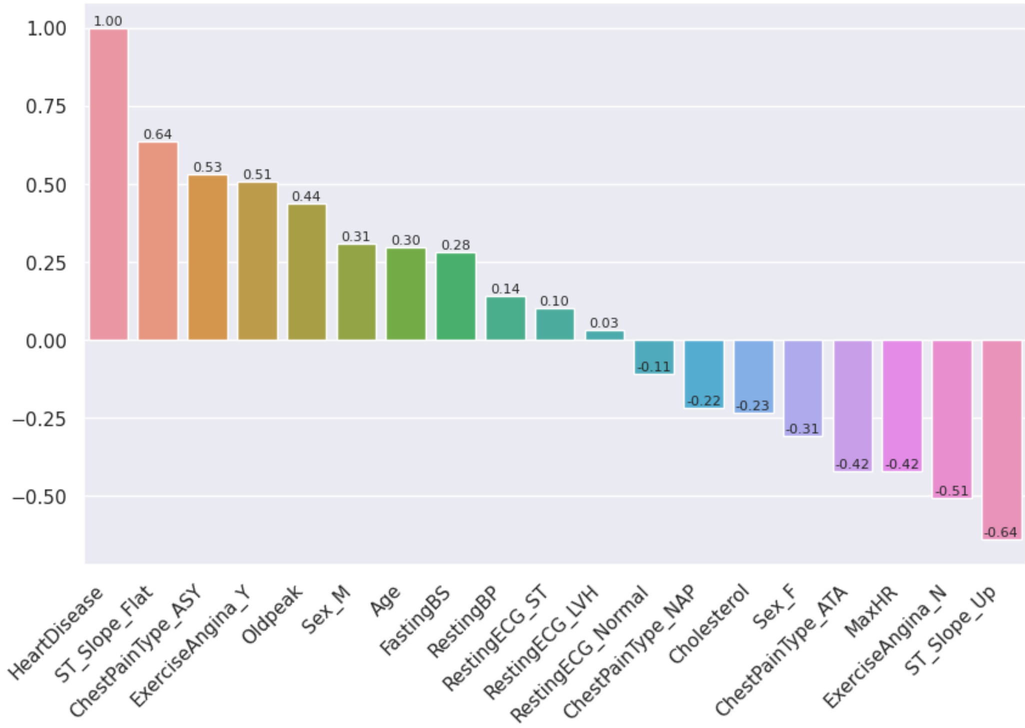

Fig. 6. Column chart with heat map of the correlation analysis of the variables in the dataset.

When the correlation values given in Figure 5 and Figure 6 are analysed, the first information that catches the eye is that gender makes a significant difference in developing the disease. With a correlation value of -0.31 for females and 0.31 for males, there is a negative relationship with the disease in females and a positive relationship in males. 'RestingBP', 'RestingECG_LVH', 'RestingECG_Normal', 'RestingECG_ST' variables have the lowest impact value.

The variables 'ChestPainType_TA', 'ST_Slope_Down', which were found to consist of fixed (same) data, were excluded from the data since they would not change the result in model training.

3.3. Classification

3.3.1. Deep neural network (DNN)

Deep learning is a sub-branch of machine learning. The main difference of deep learning methods from classical artificial neural networks is their multilayer (deep) structure. These layers consisting of interconnected neurons are divided into 3 parts. These are the input layer consisting of data features, the hidden layer consisting of more than 1 hidden unit (neuron) and finally the output layer (Figure 7). Each layer uses the output of the previous layer as input data. In this way, a hierarchical structure is created by deriving higher level features from lower level features. Deep learning has efficient algorithms for hierarchical feature extraction that represent the data and thus enables learning from the representation of the data [26].

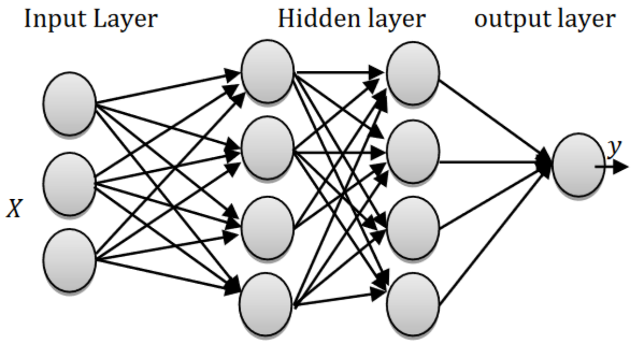

Fig. 7. Example deep neural network image.

DNN aims to minimise the error between calculated predictions and desired values by using the weights w_i of all neurons in the network. Initially randomly assigned weights are updated using the stepwise descent algorithm (Equation 3). At each iteration, a loss/error value is obtained by comparing the difference between the loss function predictions and the actual values. This loss/error is used in the gradient calculation and the weights are updated [26], [27].

$$\Delta w = -\eta \left(\frac{dE}{dw}\right) \tag{3}$$

3.3.2. Convolutional neural networks (CNN)

CNN is an artificial neural network widely used in the field of deep learning, especially in visual data processing such as image classification, object detection, handwriting recognition. There is a convolution layer for filtering and feature map extraction, a pooling layer for dimension reduction, and a fully connected layer that produces output using features obtained from previous layers. CNN is a feed-forward artificial neural network model that optimises the weights and biases of the network using the backpropagation algorithm.

CNN does not use predefined features to locally detect and learn features in the image, but instead learns locally connected neurons (filters). In other words, CNN has the ability to automatically learn features from the image [28]–[30].

3.4. Method

In this research, Deep Neural Network (DNN) and Convolutional Neural Networks (CNN) neural networks were used for early diagnosis of heart attack. For this purpose, a Deep Neural Network with 3 layers and a total of 881 parameters and a Convolutional Neural Networks network with 4257 parameters consisting of 5 layers including convolution, pooling, flatten and two fully connected layers were created. Summary information of these networks is presented in Table 3 and Table 4.

Table 3. DDN model summary.

Layer (type)	Output Shape	Param #
Dense	(None, 32)	608
Dense	(None, 8)	264
Dense	(None, 1)	9

Table 4. CNN model summary.

Layer (type)	Output Shape	Param #
Conv1D	(None, 16, 32)	128
MaxPooling1D	(None, 8, 32)	0
Flatten	(None, 16)	0
Dense	(None, 1)	4112
Dense		17

In order to find the best hyperparameters while training the models, Model Tuning (Model Validation) was performed with GridSearhCV. GridSearhCV returns a parameter array by finding the parameters with the highest success among the variables given as parameters. The parameters specific to each model were tried with experimental methods and the parameters with the highest success rate are presented in Table 5.

Table 5. Parameter space by models and best selected parameters.

Model Name	Parameter Space	Selected with GridSearhCV Best Parameters
DNN	'batch_size': [16, 32], 'epochs': [10,20,100,200], 'hidden_units': [8, 16, 32]	'batch_size': 32, 'epochs': 100, 'hidden_units': 32
CNN	'filters': [16,32, 64], 'kernel_size': [3, 5], 'epochs': [25,100,200], 'batch_size': [16,32]	'filters': 32, 'kernel_size': 3 'epochs': 100, 'batch_size': 16

In both deep learning models, 'binary_crossentropy' was chosen as the loss function, gradient-based 'adam' as the optimizer and 'accuracy' as the metrics. The loss function 'binary_crossentropy' was chosen because the loss function 'binary_crossentropy' is a successful loss function that is widely used in the literature for classification between two classes (true, false).

In addition, KFold and StratifiedKFold methods were used with CV (CrossValidation) values of 2, 3 and 5 for each model. Experimental results were obtained by dividing the data in different ratios such as 80-20, 70-30, 75-25 as training and test data. Random_state was set to a constant value for consistency (so that the data is always split at the same point for training and testing).

4. Research findings

In this study, classification models were developed with Deep Neural Network (DNN) and Convolutional Neural Network (CNN). In order to evaluate the performance of the classification models, the information obtained from Classification Report, Confusion Matrix and ROC Curve are combined and presented in Table 6. In addition, Confusion Matrix (Figure 8), LOS Curve (Figure 9) and combined ROC Curve (Figure 10) are presented in order to visually evaluate the model performance.

Table 6. Information obtained from Classification Report, Confusion Matrix, and ROC curve.

Model	Acc	AUC	Precision	Recall	F1-Score	Support	True Classification		False Classification	
							TP	TN	FP	FN
DNN	0.9375	0.96	0.9629	0.9176	0.9397	160	78	72	3	7
CNN	0.9312	0.96	0.9512	0.9176	0.9340	160	78	71	4	7

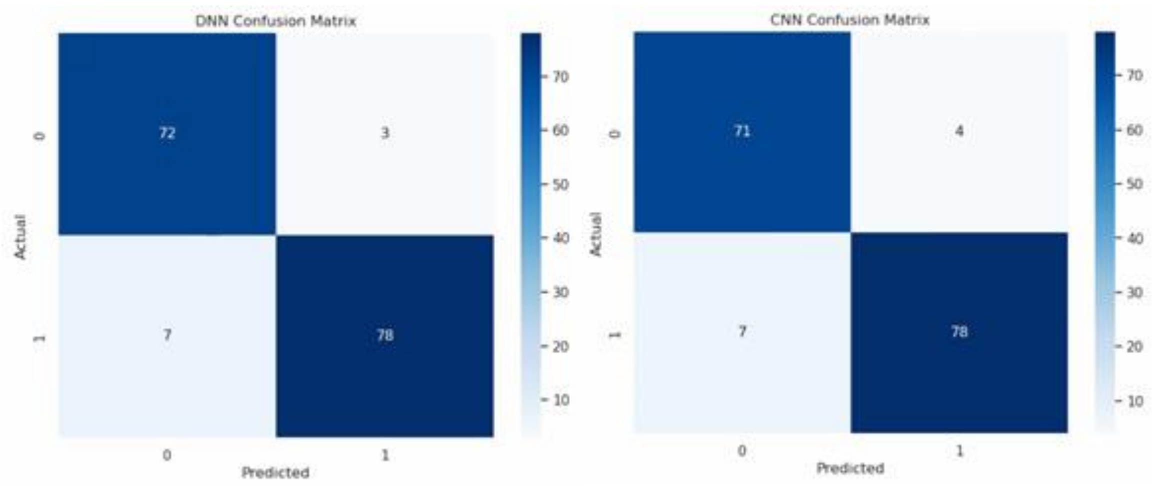

Fig. 8. DNN and CNN confusion matrix.

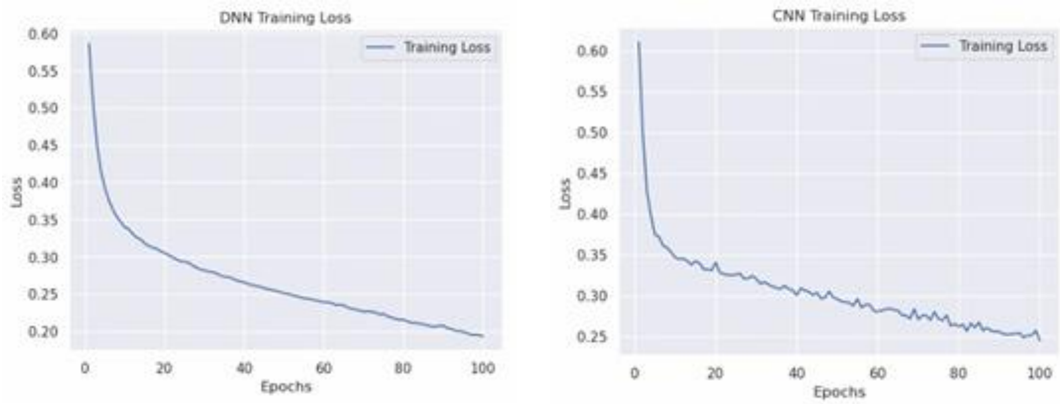

Fig. 9. DNN and CNN Loss curve.

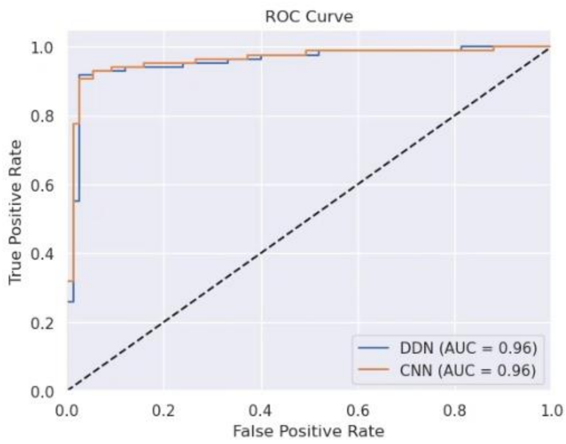

Fig. 10. DNN and CNN combined ROC curve.

5. Discussion and conclusions

In this study, a prediction model was developed with Deep Neural Network (DNN) and Convolutional Neural Network (CNN) using a heart failure prediction dataset consisting of 918 observations with the combination of 11 common features of universally accepted Cleveland (n=303), Hungarian (n=294), Swiss (n=123), Long Beach VA (n=200) and Stalog (Heart) (n=270) datasets obtained through *kaggle*. Classification Report, Confusion Matrix (Figure 8), LOS Curve (Figure 9) and ROC Curve (Figure 10), which are frequently used in the literature, were used to evaluate the performance of the classification models.

In this study, two different deep learning models were created to assist physicians in predicting heart failure and the performance of the models was compared with metrics accepted in the literature. Accuracy, Precision, Recall and F1-Score values obtained with Classification Report for the compared models were 0.9375, 0.9629, 0.9176, 0.9397 for DNN and 0.9312, 0.9512, 0.9176, 0.9340 for CNN respectively. The AUC value calculated on the ROC curve for the compared models was found to be 0.96 for both models equally. Accuracy refers to the correct classification rate of the generated model. Precision indicates how many of the instances where the generated model predicts positive are actually positive. Recall represents how many true positive samples are correctly predicted and F1-Score represents a balance of precision and sensitivity. The fact that the Accuracy, Precision, Recall and F1-Score values obtained for DNN and CNN are all between 91% and 96% shows that the model classifies the data accurately and balancedly, making very few errors and that the model works quite successfully.

The accuracy percentages of similar studies in the literature are presented in Table 7.

Table 7. Accuracy percentages of similar studies in the literatüre.

Study	Method	Sample	Accuracy (%)
(Abdullah & Rajalaxmi, 2012)	Random Forest	369	63,33
(Salman & Aksoy, 2022)	Random Forest	299	79,66
(Göktaş & Yağanoğlu, 2020)	C4.5 Decision Tree	303	83,09
(Cihan vd., 2018)	Multilayer Artificial Neural Network	303	83,49
(Shao vd., 2014)	Hibrid Model	280	83,93
(Keser & Keskin, 2023)	Artificial Neural Network	299	86,67
(Ekrem vd., 2020)	Random Forest	303	86,88
(Coşar & Deniz, 2021)	Random Forest	918	88,00
(Das vd., 2009)	Statistical Analysis System	303	89,01
(Özmen vd., 2018)	Support Vector Machine	303	89,47
(Potur & Erginel, 2021)	Multilayer Perseptron	299	90,00
(Gündoğdu, 2021)	Random Forest	303	90,20
(Çil & Güneş, 2022)	Artificial Neural Networks	303	90,54
(Coşkun & Kuncan, 2022)	Random Forest	918	90,76
(Yilmaz & Yağin, 2021)	Multilayer Neural Networks	303	91,10
(Ozcan vd., 2019)	Support Vector Machine	170	91,67
Yilmaz & Yağin, 2022)	Random Forest	909	92,90
(Vatansever vd., 2021)	Genetic Algorithm	303	93,44
This Study	**Deep Neural Network**	**918**	**93,75**

When the studies in the literature listed in Table 7 are analysed, it is understood that the success obtained in this study is higher than the previous studies.

Another important contribution of the study to the literature is that it proved that an artificial neural network such as CNN, which is widely used in visual data processing such as image classification, object detection, face recognition, etc., can also give successful results on digital data when used in 1D (dimensional).

On the other hand, in the study describes each stage from data pre-processing to model predict and evaluation in a systematic and detailed manner. It is thought that this study may contribute to researchers in similar deep learning studies that can be carried out in the future.

Acknowledgements

The authors declare that they have no conflict of interest.

References

[1] J. H. Brice, J. K. Griswell, T. R. Delbridge, and C. B. Key, “Stroke: from recognition by the public to management by emergency medical services,” *Prehosp Emerg Care*, vol. 6, no. 1, pp. 99–106, 2002, doi: 10.1080/10903120290938904.

[2] M. H. Park *et al.*, “No difference in stroke knowledge between Korean adherents to traditional and western medicine – the AGE study: an epidemiological study,” *BMC Public Health*, vol. 6, no. 1, p. 153, Jun. 2006, doi: 10.1186/1471-2458-6-153.

[3] M. H. Fazel Zarandi, A. Seifi, M. M. Ershadi, and H. Esmaeeli, “An expert system based on fuzzy bayesian network for heart disease diagnosis,” in *Fuzzy Logic in Intelligent System Design*, P. Melin, O. Castillo, J. Kacprzyk, M. Reformat, and W. Melek, Eds., in Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2018, pp. 191–201. doi: 10.1007/978-3-319-67137-6_21.

[4] H. Yan, Y. Jiang, J. Zheng, C. Peng, and Q. Li, “A multilayer perceptron-based medical decision support system for heart disease diagnosis,” *Expert Systems with Applications*, vol. 30, no. 2, pp. 272–281, Feb. 2006, doi: 10.1016/j.eswa.2005.07.022.

[5] R. Das, I. Turkoglu, and A. Sengur, “Effective diagnosis of heart disease through neural networks ensembles,” *Expert Systems with Applications*, vol. 36, no. 4, pp. 7675–7680, May 2009, doi: 10.1016/j.eswa.2008.09.013.

[6] A. S. Abdullah and R. R. Rajalaxmi, “A data mining model for predicting the coronary heart disease using random forest classifier,” *IJCA Proceedings on International Conference in Recent trends in Computational Methods, Communication and Controls (ICON3C 2012)*, vol. ICON3C, no. 3, Apr. 2012, Accessed: Jun. 12, 2023. [Online]. Available: https://www.ijcaonline.org/proceedings/icon3c/number3/6020-1021

[7] Y. E. Shao, C.-D. Hou, and C.-C. Chiu, “Hybrid intelligent modeling schemes for heart disease classification,” *Applied Soft Computing*, vol. 14, pp. 47–52, Jan. 2014, doi: 10.1016/j.asoc.2013.09.020.

[8] Ş. Cihan, B. Karabulut, G. Arslan, and G. Cihan, “Koroner arter hastalığı riskinin veri madenciliği yöntemleri ile incelenmesi,” *UMAG*, vol. 10, no. 1, Art. no. 1, 2018, doi: 10.29137/umagd.419663.

[9] Ö. Özmen, A. Khdr, and E. Avci, “Sınıflandırıcıların kalp hastalığı verileri üzerine performans karşılaştırması,” *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 30, no. 3, Art. no. 3, Sep. 2018.

[10] İ. Ozcan, B. Tasar, A. B. Tatar, and O. Yakut, “Destek vektör makinasi algoritması ile kalp hastalıklarının tahmini,” *JCS*, vol. 4, no. 2, Art. no. 2, Dec. 2019.

[11] M. E. Göktaş and M. Yağanoğlu, “Veri bilimi uygulamalarının hastalık teşhisinde kullanılması: Kalp krizi örneği,” *JISMAR*, vol. 2, no. 2, Art. no. 2, Dec. 2020.

[12] Ö. Ekrem, O. K. M. Salman, B. Aksoy, and S. A. İnan, “Yapay zekâ yöntemleri kullanılarak kalp hastalığının tespiti,” *MBTD*, vol. 8, no. 5, Art. no. 5, Dec. 2020, doi: 10.21923/jesd.824703.

[13] M. Coşar and E. Deniz, “Makine öğrenimi algoritmaları kullanarak kalp hastalıklarının tespit edilmesi,” *EJOSAT*, no. 28, Art. no. 28, Nov. 2021, doi: 10.31590/ejosat.1012986.

[14] E. A. Potur and N. Erginel, “Kalp yetmezliği hastalarının sağ kalımlarının sınıflandırma algoritmaları ile tahmin edilmesi,” *EJOSAT*, no. 24, Art. no. 24, Apr. 2021, doi: 10.31590/ejosat.902357.

[15] S. Gündoğdu, “Kalp hastalık risk tahmini için Python aracılığıyla sınıflandırıcı algoritmalarının performans değerlendirmesi,” *DEUFMD*, vol. 23, no. 69, Art. no. 69, Sep. 2021, doi: 10.21205/deufmd.2021236926.

[16] R. Yilmaz and F. H. Yağin, “A comparatıve study for the predıctıon of heart attack rısk and assocıated factors usıng mlp and rbf neural networks,” *JCS*, vol. 6, no. 2, Art. no. 2, Dec. 2021, doi: 10.52876/jcs.1001680.

[17] B. Vatansever, H. Aydin, and A. Çetinkaya, “Genetik algoritma yaklaşımıyla öznitelik seçimi kullanılarak makine öğrenmesi algoritmaları ile kalp hastalığı tahmini,” *JSTER*, vol. 2, no. 2, Art. no. 2, Dec. 2021, doi: 10.53525/jster.1005934.

[18] O. K. M. Salman and B. Aksoy, “Rasgele orman ve ikili parçacik sürü zekâsi yöntemiyle kalp yetmezliği hastaliğindaki ölüm riskinin tahminlenmesi,” *IJ3DPTDI*, vol. 6, no. 3, Art. no. 3, Dec. 2022, doi: 10.46519/ij3dptdi.982670.

[19] C. Coşkun and F. Kuncan, “Evaluation of performance of classification algorithms in prediction of heart

failure disease," *KSU J. Eng. Sci.*, vol. 25, no. 4, Art. no. 4, Dec. 2022, doi: 10.17780/ksujes.1144570.

[20] E. Çil and A. Güneş, "Makine öğrenmesi algoritmalarıyla kalp hastalıklarının tespit edilmesine yönelik performans analizi," *ABMYO Dergisi*, vol. 17, no. 65, Art. no. 65, Nov. 2022.

[21] R. Yilmaz and F. H. Yağin, "Early detection of coronary heart disease based on machine learning methods," *Med Records*, vol. 4, no. 1, Art. no. 1, Jan. 2022, doi: 10.37990/medr.1011924.

[22] S. B. Keser and K. Keskin, "Kalp yetmezliği hastalarının sağ kalım tahmini: Sınıflandırmaya dayalı makine öğrenmesi algoritmalarının bir uygulaması," *Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi*, vol. 23, no. 2, Art. no. 2, May 2023, doi: 10.35414/akufemubid.1033377.

[23] E. Cevahir, *SPSS ile Nicel Veri Analizi Rehberi Egemen Cevahir*. İstanbul: Kibele Yayınları, 2020.

[24] M. B. Wilk and R. Gnanadesikan, "Probability plotting methods for the analysis of data," *Biometrika*, vol. 55, no. 1, p. 1, Mar. 1968, doi: 10.2307/2334448.

[25] U. Çolak, "Makine Öğrenmesi - Veri Ön İşleme," Medium. Accessed: Sep. 19, 2023. [Online]. Available: https://ufukcolak.medium.com/makine-ogrenmesi-veri-on-isleme-5-58e1ce73c1fb

[26] U. A. Kimanuka and O. Buyuk, "Turkish speech recognition based on deep neural networks," *SDÜ Fen Bil Enst Der*, vol. 22, pp. 319–329, Oct. 2018.

[27] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference for Learning Representations*, 2015, doi: 10.48550/arXiv.1412.6980.

[28] F. Bozkurt, "A study on CNN based transfer learning for recognition of flower species," *EJOSAT*, no. 32, Art. no. 32, Dec. 2021, doi: 10.31590/ejosat.1039632.

[29] Y. Li, Z. Hao, and H. Lei, "Survey of convolutional neural network," *Journal of Computer Applications*, vol. 36, no. 9, pp. 2508–2515, 2016.

[30] Z. Li, W. Yang, S. Peng, and F. Liu, "A survey of convolutional neural networks: Analysis, applications, and prospects," 2020, doi: 10.48550/ARXIV.2004.02806.