Health and Technology
https://doi.org/10.1007/s12553-020-00509-3

ORIGINAL PAPER

# Multilayer perceptron based deep neural network for early detection of coronary heart disease

Nancy Masih[1] · Huma Naz[1] · Sachin Ahuja[1]

Received: 27 July 2020 / Accepted: 5 November 2020


**Abstract**
Coronary heart disease leads to a high mortality rate worldwide. Owing to delays in its detection, its treatment becomes challenging with little chances of recovery in many cases. An efficient, early-stage detection method is therefore urgently needed. Using the Framingham Heart Study Dataset, this study shows how data pre-processing via the multilayer perceptron following a deep learning approach will improve data quality when computing the likelihood of one having coronary heart disease. Apart from being highly efficient, our proposed approach results in highaccuracy of 96.50%. Finally, the paper discusses the rise in efficiency and accuracy achieved via use of deep learning techniques to enhance predictive outcomes v. traditional ones. The proposed study attempts to detect Coronary Heart Disease at an early stage.

**Keywords** Multilayer perceptron · Artificial neural network · Classification algorithms · Data pre-processing · Coronary heart disease

## 1 Introduction

Among numerous critical diseases, Coronary Heart Disease (CHD) seems to be responsible for the high mortality rate [1] This particular disease needs immediate attention. Also, there are other heart diseases like coronary artery disease which can lead to heart attacks [12]. The major reason for deaths in developed countries is CHD. Smoking, heredity, high cholesterol, obesity, and diabetes are considered as the risk factors which are responsible for leading heart diseases [33]. Some people think that growing age is also one of the factors responsible for heart disease. But, there is no such constraint applied on age factor for heart disease as an infant can also have heart disease depending upon a deficiency of vitamin D [41].

A symptom is an evidence that shows the conditions are inappropriate for a normal human being. Therefore, an immediate action is needed to cure a before a serious and major threat to patient life. Symptoms play a vital role not only in early detection of the disease but it can improve the treatment, medical care, and life quality of patients in comparison to the patients with late-diagnosed [12]. Likewise, heart disease symptoms include anxiety, chest pain, and fatigue,it is because the heart is unable to function properly. Various studies show that chest pain is the most responsible factor for leading heart disease [17]. However, [10] claims that chest pain and anxiety can be considered as the consequences of heart disease rather than symptoms. Chest pain increases depression and anxiety in a person which is also a serious issue [10]. Treatment becomes quite easy if the disease is identified at an initial stage. Diagnosis of heart disease is possible with the medical knowledge collected from patients, yet a true examination of the disease cannot be achieved sometimes due to various issues perhaps. There are different conventional approaches available that are used for the treatment of heart disease. But, those approaches are vulnerable to human errors because different doctors and clinicians vary in their knowledge and experiences, whereas machine learning techniques tends to be proved more efficient to process raw data and transform it into useful patterns (Son, Kim, Kim, Park, & Kim, 2012). To overcome the numerous failures of

✉ Nancy Masih
nancy@chitkara.edu.in

Huma Naz
huma.naz@chitkara.edu.in

Sachin Ahuja
sachin.ahuja@chitkara.edu.in

1 Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Published online: 10 November 2020

conventional or traditional methods used for the prediction of CHD, machine learning techniques can come into the practice for detection of CHD prevalence. Machine learning is giving an excellent performance in various aspects of new computing technologies. Machine learning models can predict independently after the training process. Machine learning techniques can also be used for prediction. A number of diseases can also be predicted using classification algorithms depending upon the requirements. Prediction can be done by recognizing all the influencing attributes and their associated factors present in the raw data [37]. A continued effort has been made in the field of CHD which involves prediction, analysis, and treatment using data mining techniques and it achieves great accuracy [30]. Classification algorithms such as Decision Tree, Logistic Regression, Naive Bayes, Support Vector Machine, Artificial Neural Network, and Deep Learning can be used to determine the target class. There are distinct application areas where deep learning algorithms are used for prediction. Deep learning is a subset of machine learning which in turn is a subset of artificial intelligence that has the capability of learning through unsupervised data. Deep learning is a renowned machine learning algorithm inspired by the structure of the human brain [24]. Deep learning makes use of a computational model that consists of multiple processing layers to work on multiple layers of abstraction for data representation learning. In recent years, deep learning achieved huge success in improving results' accuracy in various domains like visual object recognition, speech recognition, object detection, face detection, and prediction [56, 57]. In deep learning, a multi-layer perceptron is the class of deep neural networks. Neural networks or multi-layer perceptrons in the field of artificial neural network which follows the approach of deep learning. However, a single layer perceptron is a single neuron outcome to larger neural networks [27].

This discipline inspects how the biological brain can be used to design predictive models which can accomplish difficult tasks. The target of using deep learning is not to create a brain model but to create a model that follows approach of human brains' neurons [29]. The potential of neural networks comes from the ability to learn and relate the data to the predicted output class. Neural networks are capable to learn any mapping function and prove to be a universal approximation algorithm. Because of the efficient outcome of Multilayer Perceptron (MLP) in difficult to complex tasks, it is commonly being implemented for prediction of CHD.

The paper aims to accomplish two major goals: The first is to find a predictive model using the MLP approach to predict CHD disease and the other one is to conduct a comparison study with other existing models.

## 1.1 Data description

For Coronary heart study, the Framingham Heart Study (FHS) dataset has been used. FHS dataset is dedicated to the people living in Framingham. To predict the risk factor, FHS dataset has been used. The dataset consist of 18 attributes with 4583 instances. A huge number of scientific papers have been published using the FHS dataset. As there are 18 attributes in the dataset, the most influencing attributes are being recognized using pre-processing techniques to predict heart disease.

## 1.2 Data pre-processing

Pre-processing is the method applied to the dataset before its processing. Typically, pre-processing modify the raw data which can enhance the classification ability of processing. Also, it includes detection and removal of outliers where outliers are the inconsistent type of data. It also performs attributes extraction, normalization, integration, aggregation, and discretization [21]. The dataset consists of 18 attributes with 4583 total number of instances, in which the following 6 attributes only as shown in the Table1, are considered as the most prominent attributes in the dataset.

However, it is being analyzed that the dataset contains several missing and inconsistent values, and those values were processed after applying data pre-processing techniques on it to improve the quality of the data. Therefore, Data cleaning operation has been performed on the dataset to remove the noise and integrate the various attributes. Moreover, the dataset consists of some missing values which need to be eliminated to attain the information as shown in the following Fig. 1.

The process of conversion of raw data into an applicable manner for effective analysis is known as data pre-processing, which includes:

- collection of data from various sources
- removal of noise and duplicate entries in data
- selection of most influencing attributes

## 1.3 Steps involved in data preprocessing

### 1.3.1 Need for data preprocessing

Datasets can contain numerical, nominal, or polynomial values. But different algorithms can process only a particular type of data. So there is a need to process the data to transform the data into the required form of data type which can be done only by using data preprocessing

**Table 1** The most prominent attributes from FHS dataset are extracted using data preprocessing techniques

| | |
|---|---|
| **Blood Pressure** | The normal range of systolic and diastolic BP (Blood Pressure) is less than 120 and 80 respectively. The range if reaches higher than this scale, then it can be responsible for the case of coronary arteries disease [34] |
| **Gender** | The disease is not biased for any gender. It can occur in men as well as women. It is being observed that both male and female gender is vulnerable to CHD equally. However, according to Paynter et al. although CVD takes 8–10 years to develop in women, but still it is recognized as one of the major causes of death in women (Paynter et al., 2018a) |
| **Age** | Following the study of Gordon et al. it is concluded that people with the age of 65 or more have high chances of having heart disease. Because the heart begins to weaken in the growing age [14] |
| **Total-Blood Cholesterol** | High as well as low blood pressure is not appropriate for a person. But in context to CHD, high cholesterol is riskier [3] |
| **Body Mass Index** | BMI is the short term used for Body Mass Index. It is the value derived from a person's weight and height [38]. Even, National Institutes of Health (NIH) defines the health problems like Obesity and overweight using BMI formula which is: $BMI = \frac{\text{Weight (Kg)}}{(\text{Height (m)})^2}$ |
| **Smoking** | Smoking is one of the strong aspects which tend to be responsible for Coronary heart disease. It can damage the heart arteries and increase the rate of CHD development due to fats accumulation in the arteries. The probability of this particular disease is directly proportional to the number of cigarettes an individual smokes per day. Therefore, to wrap it all, we can say that smokers have higher chances of having heart disease as compared to non-smokers. According to Birgit-Christiane Zyriax, there are very few women who smoke and smoking seems to be more hazardous for them in comparison to men [56, 57]. Women smokers are 25% more vulnerable than men to catch this disease [43] |
| **Death from CVD** | This attribute is considered as the class or label for the classification. The label is categorized as "Yes" or "No" accordingly |

techniques. Attributes with different variety of values are overestimated in the process of selection of most influential attributes, for deriving the decision rules and ending with a decision as well. Because, only a few sets of attributes are responsible for achieving the target class [7]. There may be redundant values also which is needed to eliminate for improving the data quality. Apart from attribute selection, a very prominent task to reduce the data size has also been performed which is commonly known as data reduction to improve the efficiency of the data [25]. Because the dataset may have a large number of attributes that seem to be redundant or duplicate is needed to be eliminated using data preprocessing techniques. Different steps involved in data pre-processing is shown in Fig. 2.

### 1.4 Data cleaning

**In case of missing data:** Sometimes, data collected can contain missing information which can be handled by using the following methods:

- **Eliminate missing values:** Handling missing values is one of the steps performed in data preprocessing. Missing values can be eliminated by using global values, by eliminating the whole row with missing value, by applying average mean value [13]. In the FHS dataset, numbers of instances were identified to be missing, then all the missing values were eliminated by deleting the whole row.

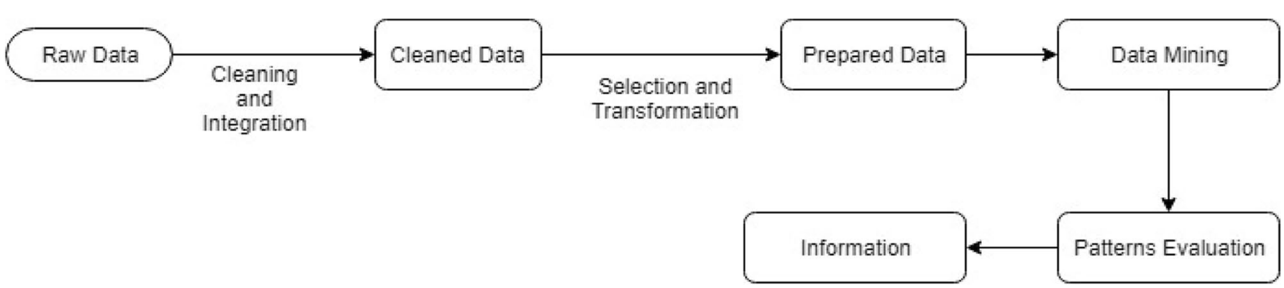


**Fig. 1** Targeted information achieved after undergoing the data pre-processing stage

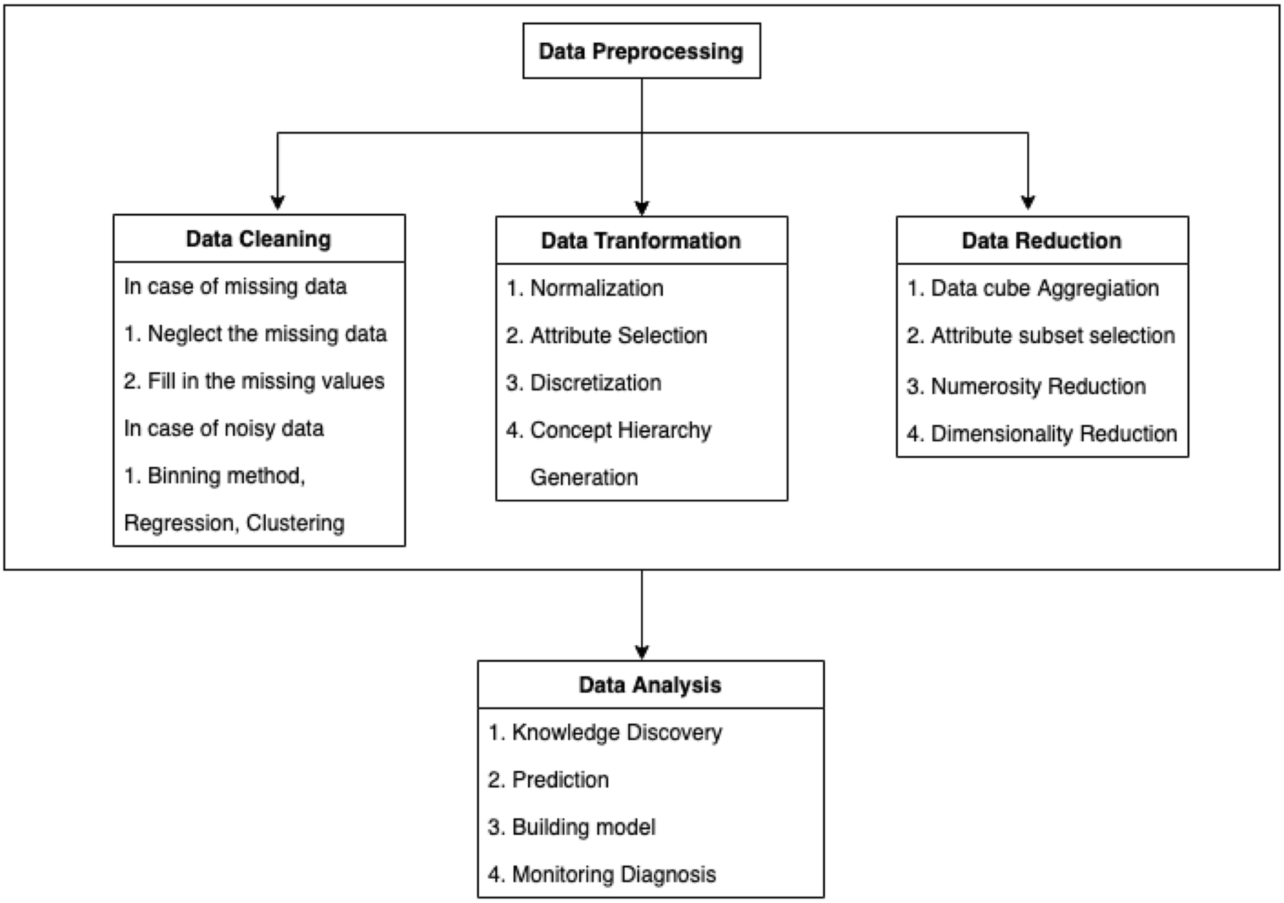


**Fig. 2** Different steps in Data Preprocessing ("Data Preprocessing in Data Mining," [9])

- **Remove noisy data:** A random error or meaningless data encounters in the dataset is known as Noise [13]. Likewise, some data was recognized to be inconsistent or consisting outliers which have been removed and using the suitable consistent values. For example, values in blood pressure attribute were integer values which were being converted into string values.

## 2 Related work

The importance of Data mining and its techniques has been well-recognized in literature. Numerous data mining techniques can be used for distinct intents. Data mining is a multidisciplinary field which can be used in other areas applications such as fraud detection [57], consumer behavior prediction [19], medical treatment [26]. and analyzing criminal patterns [56]. Data mining techniques can be categorized into various types such as Classification, Clustering, Regression, Outlier detection, Sequential patterns, Prediction, Association rules, etc. The focus of this paper is to predict Coronary heart disease using prediction data mining techniques. So, the basic classification or prediction techniques which are available for prediction purpose are named as Decision tree, Naive Bayes, SVM, regression, and neural networks. We attempt to implement deep learning- MLP neural network model on FHS dataset. A standard Neural Network consists of multiple processors known as neurons of the neural network, each neuron produces a sequence of real-value activations. Input neurons get activated through sensors perceiving from the environment, while other neurons get activated through weighted connections from previously activated neurons [45]. However, MLP is a class of feed-forward artificial neural networks. MLP consist of an input layer, multiple hidden layers, and an output layer. Training of the model involves adjusting the weights to minimize the error rate of the model [51]. Various types of techniques can be used for prediction which produces desired results almost, but deep learning tends to be more suitable than standard data mining techniques because deep learning brings multiple advantages in learning multiple levels of representation of natural language [59]. Moreover, deep learning technique has been used for prediction of improved protein structure.

Protein is the fundamental process of Biology, and the function of protein solely relies on the structure of amino acids. Hence, protein structure prediction is very important to predict the function of the protein. Although, number of experiments have been conducted to determine the structure and they have achieved efficient results but they lack ease and time efficiency. Therefore, [46] proposed a deep learning approach for protein structure prediction [46]. [42] made use of deep learning on retinal fundus photographs to predict cardiovascular risk factors, they show that deep learning can extract new knowledge from a dataset consisting of retinal fundus images. They have also shown that the trained deep learning models can produce anatomical features to generate each prediction [42]. Parkinson's disease is a disease which leads to difficulty in talking, walking, balancing, and coordination. This disease begins gradually and gets worse with the passing time. This disease cannot be cured but it can be controlled with proper medication and treatment. Many studies have been carried out to analyze the risk factors and to find a better treatment for this disease. Deep learning is one of the algorithms used in the research area of Parkinson. [16] proposed a method to predict the severity of Parkinson's disease on the dataset of UCI's Parkinson's Telemonitoring Voice Data Set of patients using deep learning. They have used a TensorFlow deep learning library for generating a prediction model with 16 biomedical voice features and Motor UPDRS score as the target variable [16]. The conducted experiment achieved an accuracy of 83.367% and 81.6667% for train and test dataset respectively which is 37.36% greater than the accuracy produced by [35]. The stock market is the market where investors buy and sell the shares and made a profit. Investors can make a huge profit from the shares in the stock market. The stock market always reserves a significant position in the business world. Ritika Singh and Shashi Srivastava used deep learning for stock market prediction. They stated that an artificial neural network fails to provide any efficient results on the prediction of the stock market. So they introduced a deep learning approach for prediction on Google stock price multimedia data (chart) from NASDAQ [48]. According to [28], numerous approaches implemented on transportation network analysis seem to have limitations due to cumbersome data and unrealistic assumptions. However, with the growth of Intelligent Transportation Systems (ITS) and the Internet of Things (IoT), data transportation becomes very important. Therefore, [28] extended a deep learning approach for data traffic analysis. As expected, deep learning turns out to be an efficient technique for analysis of high dimensional data. Moreover, the study conducted a comparison of different algorithms on traffic congestion analysis at 60 min level of aggregation [60]. Over the last few decades, deep learning is considered one of most the powerful approaches to handle a large amount of data.

Designing a controller that can provide appropriate performance for driving scenarios autonomous is a very challenging task because the driving scenario in a simulated environment is very difficult to handle which may produce diverse errors. The number of methods fails to provide such a bug-free environment. However, deep learning methods tend to be a promising approach for providing efficient performance for complex and non-linear control problems as well [22]. Deep learning is a very promising approach for extracting significant information from complex data with a high level of abstractions. Such kind of techniques represents the data in a hierarchical manner where high-level abstract features can be modeled as low-level features. Because of this, deep learning appears to be quite beneficial in the case of unsupervised and complex data. Deep learning tends to provide invariant data representations and improved classification results [53]. A hybrid convolution neural network has recently achieved efficient results because deep learning can work with a high level of complexity present in data. Due to improved results achieved by deep learning approaches, [61] proposed a method known as adaptive windows multiple deep residual networks for speech recognition. The proposed model attempts to increase the robustness of the model by reducing the recognition error rate. The proposed system can be used as a vocal input for various expert devices. The obtained results demonstrate that the designed system reduces the recognition error rate by 7% as compared to methods of the state-of-the-art in some speech recognition tasks. Hence, the presented study of the paper reveals that deep learning can be considered as the best tool for classification purposes as it not only recognizes speech but can improve the task with minimal error rate. However, the designed system can be used in other application areas also such as EEG (electroencephalogram), ECG (Electrocardiogram), TTS (text to speech), and image recognition [61]. Varun Sapra and Madan Lal Saini conducted a study on the dataset collected from a medical college under the guidance of doctors. They reduced the features of the dataset by selecting the most prominent features using data preprocessing techniques. The deep learning model classifies the data very accurately with an accuracy rate of 97.41%. Keras library has been used for the construction of a multilayer deep learning model [44]. Congenital heart disease (CHD) is one of the reasons responsible for the mortality rate of almost 1% of live births. Infants who are suspected of this heart disease need immediate surgery for survival. To detect the disease the CMR (Cardiovascular MR) is being used but it needs high manual annotations also which should be eliminated to reduce the human error rate. Therefore, a semi-automatic or fully-automatic segmentation method is required to reduce the time, effort, and workload of clinicians. CNN (Convolution neural network) model is being designed for the detection of heart disease. The proposed model was able to segment the myocardium and the blood pool without any expert's intervention [58]. Authors presented work on recognition of core 35 Gurumukhi characters using CNN and achieved an accuracy of 98.32% which is very high as compared to previously conducted experiments [20]. [23] proposed an ensemble deep learning model which includes CNMP (Coding Network with Multilayer Perceptron) and DCNN (Deep Convolution Network) to overcome the limitations of model layers and channels due to high resolution of medical images. The researchers achieved great accuracy of 90.1% and 90.2% on medical image classification on two different datasets of HIS2828 and ISIC2017 respectively [23]. Wang et al. presented a pathological brain detection system for Alzheimer's disease (AD) detection which made use of three significant components known as wavelet entropy, multilayer perceptron, and biogeography-base optimization for early detection of the disease. However, the comparative study of the system suggests that the proposed approach tends to be the most accurate among all the latest 6 approaches used for Alzheimer's disease detection [54]. [6] presented a novel approach for emotion classification using MLP and Kohonen Self organizing map (KSOM) classifiers. The process follows three steps to detect the emotion of an image which includes face detection, features extraction, motion estimation, etc. The obtained results show an improvement in emotion detection using a multilayer layer perceptron classifier [6]. Natural language processing is one of the important domains in machine learning. When a machine can understand and analyze human natural language then the process is known as Natural language processing. Singh et al. attempt to process natural language sentences using a MLP model. The model assigns syntactical categories to each word of the sentence and transforms the sentences into patterns that can be understandable by machines [47]. [8] have made use of MLP neural network to detect saltwater anglers in marine environments. The study provides insights on threats concerning marine environmental [8]. [5] presented a comparative study of classification algorithms on various medical datasets which include Chronic Kidney Disease, Breast cancer, Cryotherapy, Hepatitis, Immunotherapy, Indian liver patient dataset, Liver disorders, Pima Diabetes, Risk factor cervical cancer, and Starlog dataset. Among all the datasets, the MLP has achieved the highest accuracy in the case of Chronic kidney disease [5]. Numerous studies have been conducted to predict the occurrence of any disease using MLP as mentioned in the literature review. But it can be seen in all the previous studies that they don't have used data preprocessing techniques which can improve the accuracy of the outcomes. However, the paper attempts to predict CHD prediction using data preprocessing techniques, which proves to be more accurate as compared to the results without using Data Preprocessing techniques.

## 3 Advantages of MLP over other techniques

There are several learning techniques available for prediction. But MLP is more efficient as compared to other classification algorithms. The major advantages of the MLP algorithm include dealing efficiently with difficult to complex problems. Furthermore, multiple applications in the area of Artificial Neural Network have been explored in the literature review which confirms the evidence of the generality and capability of the MLP learning scheme.

Therefore, it has much better learning efficiency than other classification algorithms. The number of experiments held on different widely used classification data sets exerts that the proposed algorithm achieves better results than the existing state-of-the-art learning techniques.

## 4 Methodology and model design

### 4.1 Deep feedforward neural networks

Deep feedforward neural networks are commonly called as MLP or feed-forward neural network. A neural network originates from a very famous machine learning algorithm known as perceptron [31]. Perceptron is a linear classifier with a mapping function that partitions feature space using a linear function which is a boundary line used to classify the data into 2 classes [39]. MLP neural network is a model with multiple layers of input that are associated with some weight which are presented in the processor in **a feed-forward manner** [50] as shown in Fig. 3.

The Fig. 3 demonstrates the architecture of the feed-forward network where the input layer is denoted as $X_1$ and 1 is the bias value, whereas hidden layers are associated with the activation function represented as F1, F2, and F3 as shown in the figure. Each input neuron whether travelling towards the hidden layer or output layer has some weight connected to it. The processor uses an activation function to produce the output which is represented as F1, F2, and F3 in the hidden layer's nodes [15]. If the predicted output in the output layer is the same as the desired output, then the performance is considered satisfactory, no change will be made in the weights otherwise weights will be updated to reduce the errors.

### 4.2 Why the model is known as the feed forward network

A neural network is a network of interconnected neurons known as nodes of the system. It is an attempt to make a computational model of the human brain which performs computing operations like the neural system of the brain. The main motive is to develop a system to perform computational tasks more efficiently than conventional systems.

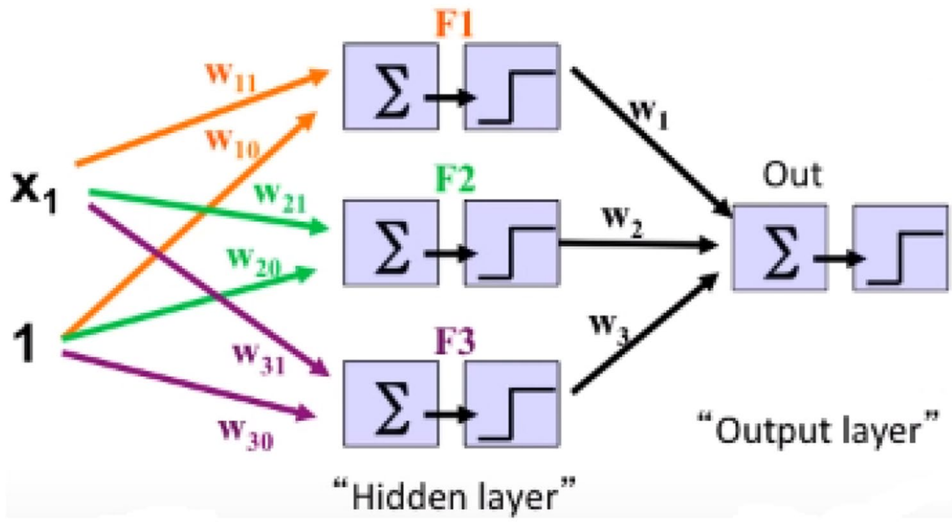


**Fig. 3** Multilayer Perceptron model

The model is known as the feed-forward model because the data can move only in the forward direction through input and hidden layers, it cannot be backpropagated. The model consists an input layer and an output layer and millions of hidden layers in between, the accuracy of the model is directly proportional to the number of hidden layers which means more the hidden layers a model consists, more the accuracy will be improved [28]. Figure 4 illustrates the flow of information travelling only in one direction which is towards an output layer**.**

## 5 Methodology and techniques

The authors have employed the MLP followed by data preprocessing methods to predict CHD. Figure 5 shows the major steps that are involved to predict CHD disease using.

### 5.1 Algorithm of the technique

1. The first and foremost step is to input the FHS dataset. FHS comprises of 18 attributes and 4583 instances of different patients.

2. As FHS dataset is a huge dataset there exist inconsistencies in data such as noise, outliers, and missing values. To eliminate the inconsistencies from the data, the authors have applied the equal width binning method to divide the attributes with inconsistencies into equal groups of bins using Eq. 1.

$$G_N = \frac{(\max(A) - \min(A))}{N} \tag{1}$$

Where $G_N$, represents the groups of bins. $A$ represents the attribute that has inconsistency and N represents the number of groups of equal bins. Then mean binning method has been applied to each group of bins to fill the missing values using Eq. 2.

$$M_{G_N} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{2}$$

Where $M_{G_N}$ used to compute the mean for each bin and n is the number of values in one bin. After filling the missing values, it has been analyzed if there is an outlier in the

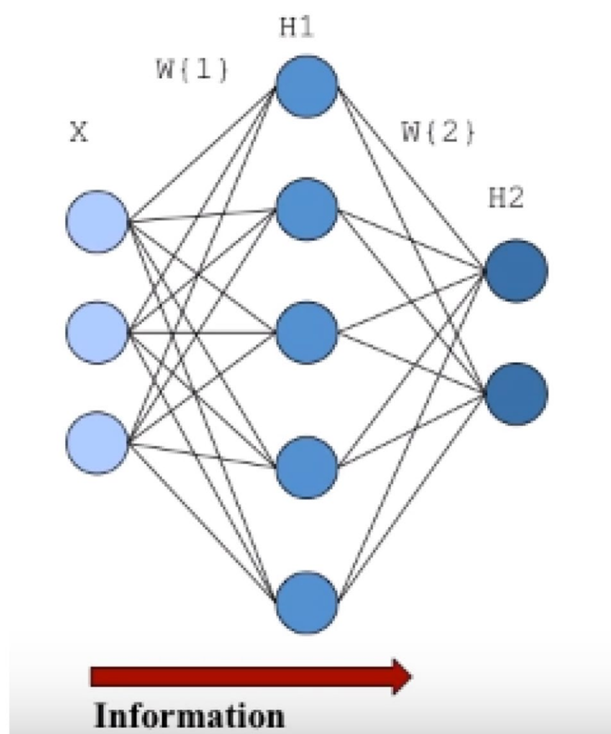


**Fig. 4** Feed Forward Network

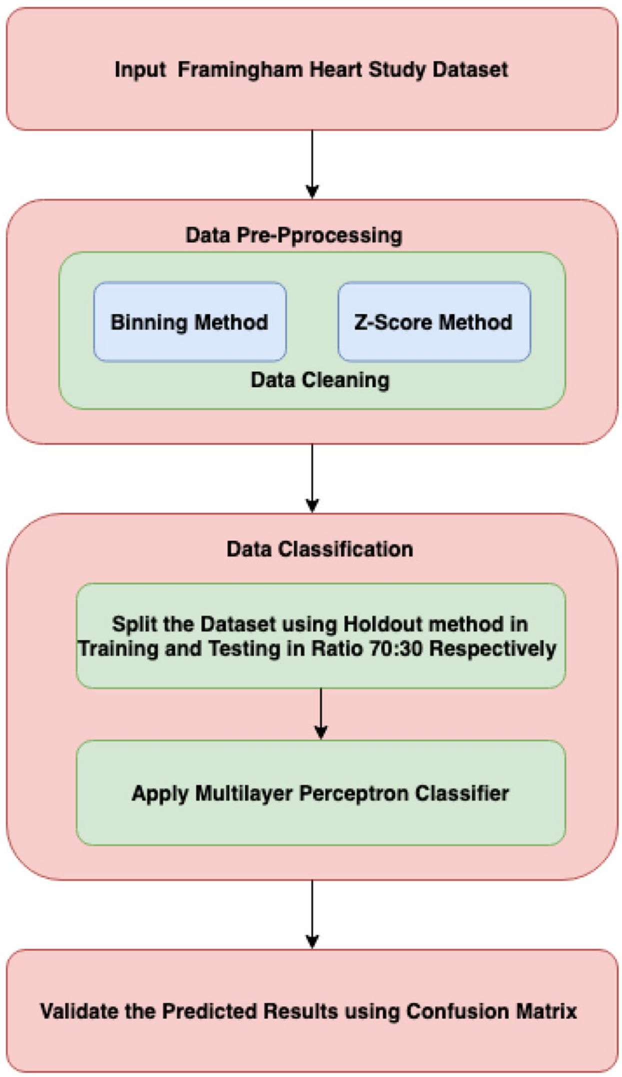


**Fig. 5** Steps followed to predict CHD

dataset. To remove the detected outlier Z-score method has been employed using Eq. 3.

$$Z_S = \frac{O - M}{S_D} \tag{3}$$

where $Z_S$ is the Z-score computed to replace the outlier O from the attribute A. M represents the mean value computed for attribute A that consists of outliers. $S_D$ is the standard deviation for the respective attribute A and can be computed using Eq. 4.

$$S_D = \sqrt{\frac{\sum (X_i - M)^2}{n}} \tag{4}$$

3. The next step is to split the pre-processed data using the holdout method into training and testing in 70:30 ratios respectively. Apply the multilayer-perceptron classifier as discussed in the previous section on the pre-processed dataset to predict the CHD. Validate the trained model against the test dataset.

4. Test and validate the predicted outcomes using evaluation matrices such as accuracy, sensitivity, and specificity.

**Table 2** Performance of MLP before data pre-processing in the confusion matrix

| | Predicted No | Predicted Yes | |
|---|---|---|---|
| Actual No | 1275 | 97 | 1372 |
| Actual Yes | 1 | 1 | 2 |
| | 1276 | 98 | |

## 6 Experimental results and discussion

This section overviews the outcomes achieved using MLP neural network before and after applying data pre-processing techniques. A multi-layer perceptron can work with a large dataset. Therefore, proposed work has been done using multiple parameters like batch size = 128, epochs = 128, and dropout = 0.6 which is used to reduce overfitting. The functions which have been used in the experiment are: Loss function = mean squared error, learning rate = 0.001, stochastic gradient descent method for an iterative process to optimize the model. The experiments are conducted using Python language, for which Anaconda environment setup had done. To implement the proposed model NVidia GeForce GTX GPU, i5 processor, window 10 is used for achieving high performance. Model designing that provide anaconda navigator and shell for backend environment setup and for multiple important supportive libraries installation which includes Numpy library, used for image array computation. Keras, an open-source neural network library which runs on top of TensorFlow has been used to design CNN model and hidden layers are created with the help of transfer function and activation function. Tensorflow provides high level and low-level API for making Keras user-friendly.

The proposed model of MLP shows an accuracy of 96.50% which is very high using the FHS dataset. Along

**Table 3** Performance of MLP before data pre-processing in the confusion matrix

| | Predicted No | Predicted Yes | |
|---|---|---|---|
| Actual No | 974 | 17 | 991 |
| Actual Yes | 31 | 352 | 383 |
| | 1005 | 369 | |

**Table 4** Accuracy, Sensitivity, and Specificity

| Using MLP | Accuracy | Sensitivity | Specificity |
|---|---|---|---|
| Before Pre-processing | 92.87% | 50% | 92% |
| After Pre-processing | 96.50 | 91.90% | 98.28% |

with this, the paper describes the rise in efficiency achieved in the model after using Data preprocessing techniques. At last, a comparison has also been conducted between the prediction outcome achieved with or without data preprocessing techniques as demonstrated in Table 2 and Table 3.

**Evaluation metrics:** Different evaluation metrics are used to compute the essence of the proposed prediction model. Evaluation metrics tend to play a significant role in measuring the performance of the prediction model. There can be a case where model can perform well with a specific evaluation metric but it can be considered as less accurate when any other evaluation metric comes in use. So, the selection of suitable performance metrics is an essential part of the development. Several evaluation metrics are available for performance evaluation of classification models like Classification Accuracy, Sensitivity–Specificity, F1-score, ROC (receiver operating characteristic curve), AUC (area under the curve) which can be used to measure performance. But the evaluation parameters which are considered to evaluate the performance of the applied technique considered are Accuracy, Sensitivity, and Specificity which are discussed below in detail.

1 **Accuracy:** Accuracy of a measurement is ratio of correctly classified observations to the total number of observations. The goal of the study is to classify an individual as "Diseased" or "Non-diseased". So, the accuracy will be based on the rate of correctly classified the labels of "Diseased" or "Non-diseased" [18].
2 **Sensitivity:** It is a proportion of true positive correctly classified. In the context of the proposed study, it can be defined as how often we correctly diagnose the CHD' individual as having a disease [52].
3 **Specificity:** It is a proportion of true negative correctly classified. In the context of the proposed study, it can be defined as how often we correctly diagnose a non-disease individual as not having a disease [36].

Table 4 demonstrates Accuracy, Sensitivity, and Specificity rate of the results obtained using a MLP before Data pre-processing and after Data Pre-processing respectively. It also shows that sensitivity and specificity eventually increase with the increase in accuracy rate.

Figure 6 shown above shows the graphical representation of the variation comes in accuracy after pre-processing the FHS dataset using Data pre-processing techniques. Improvement of 3.60% has been observed using data pre-processing techniques.

## 6.1 Comparison of the proposed technique with other learning techniques

Authors have used the same set of input datasets generated by the FHS. The results are summarised in the following table using three evaluation metrics named Accuracy, Sensitivity, and Specificity. The proposed work provides more accurate results as compared to the latest studies conducted using FHS Dataset as indicated in Table 5.

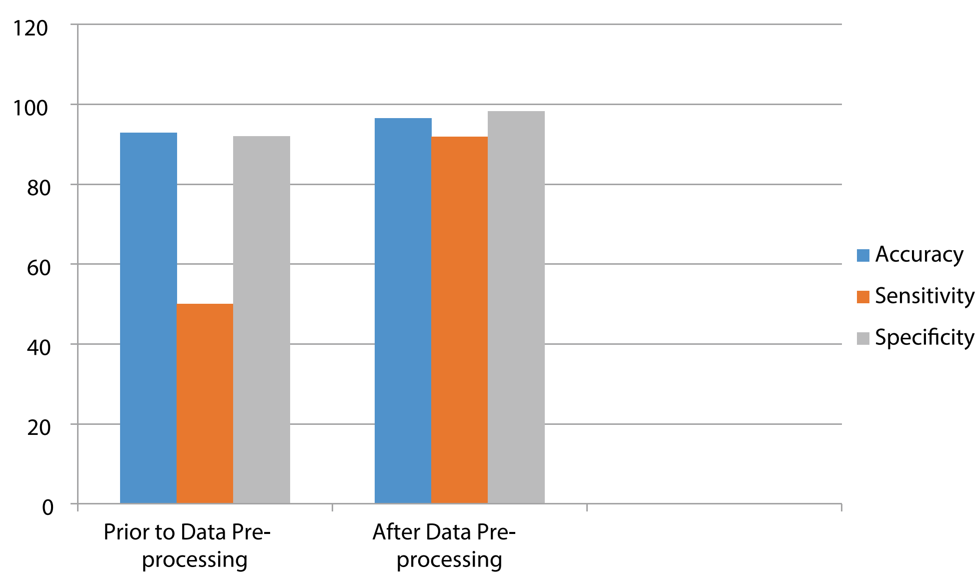


**Fig. 6** Bar graph shows variations in results before and after the pre-processing of data

**Table 5** Comparison of the proposed technique with previous studies

| Dataset | Year | Author | Technique | Accuracy % | Sensitivity % | Specificity % |
|---|---|---|---|---|---|---|
| FHS | 2015 | [4] | MLP | 92.10 | 50 | 93 |
| FHS | 2018 | [11] | Random Forest | 78 | 75 | 80 |
| FHS | 2020 | [2] | Voting ensemble classifier | 73 | 72 | 77.45 |
| FHS | | | **Proposed technique** | **96.5** | **91.9** | **98.28** |

## 7 Conclusion and its future scope

The proposed study presented a medical decision support system based on MLP deep neural network to predict CHD. In particular, FHS dataset has identified several inconsistent values which are being improved using Data pre-processing techniques. Moreover, 7 most crucial attributes were selected for the diagnosis of the heart disease and encoded them accordingly. Then, the model is trained by using 70% data and 30% has been used for test purposes and then the model is being designed using a MLP deep neural network which attains 96.50% accuracy. Three assessment metrics have been used which are Accuracy, Sensitivity, and Specificity to measure the performance of the proposed system. The results show that the proposed system has attained a very efficient prediction outcome. Besides this, the improvement in the quality of data has also been demonstrated using data pre-processing techniques which shows 3.63% of improvement in the prediction accuracy.

## 8 Future scope

This research work has assertive constraints that must be in consideration. The research has been proposed on the FHS dataset which is available on the UCI repository. Also, this dataset has approximately 4583 number of instances. The more the data is comprehensive, the more the results are convincing. Eventually, the MLP produces enough accurate results for the prediction of CHD. However, clustering techniques can also be used to attain a hybrid approach of a MLP with clustering. The last area of improvement is the methodology of feature extraction and construction. Though a limited feature construction has been done in the proposed work, it results in the collaboration of the most convincing attributes in the analysis. By using a more comprehensive data set, multiple tables can be extended. Therefore, more custom variables can be developed using those extended tables. But there is a limitation of developing more custom variables because it will become more difficult to define the relationships between it and the variables relying on it.

FHS Dataset has been taken for this particular study which is freely available on this link**:** https://www.kaggle.com/naveengowda16/logistic-regression-heart-disease-prediction

## References


1. Mackay J, Mensah GA. The atlas of heart disease and stroke. World Health Organization; 2004.
2. Alfred R, Yuto L, Haviluddin H, Chin KO, eds. Computational Science and Technology: 6th ICCST 2019, Kota Kinabalu, Malaysia, Springer Nature. August 2019;603:29-30.
3. Arsenault BJ, Rana JS, Stroes ESG, Després J-P, Shah PK, Kastelein JJP, Khaw K-T. Beyond low-density lipoprotein cholesterol. J Am Coll Cardiol. 2009;55(1):35–41. https://doi.org/10.1016/j.jacc.2009.07.057.
4. Hashim A, Bakhteri R, Hau YW. "Arrhythmia detection based on Hermite polynomial expansion and multilayer perceptron on system- on-chip implementation." ARPN J Engr Appl Sci. 2015;10:20.
5. Bendi VR, Boddu RSK. Performance Comparison of Classification Algorithms on Medical Datasets. EasyChair; 2020; No. 2322.
6. Berkane M, Belhouchette K, Belhadef H. Emotion recognition approach using multilayer perceptron network and motion estimation. International Journal of Synthetic Emotions. 2019;10(1):38–53.
7. Chandrasekar, P., Qian, K., Shahriar, H., & Bhattacharya, P. Improving the prediction accuracy of decision tree mining with data preprocessing. In 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC). 2017;2:481-484. IEEE.
8. Chi YN, Chi J. Saltwater anglers toward marine environmental threats using multilayer perceptron neural network framework. International Journal of Data Science and Advanced Analytics. 2020;2(2):6–17.
9. Data Pre-processing in Data Mining. In: GeeksforGeeks. 2019. https://www.geeksforgeeks.org/data-preprocessing-in-data-mining. Accessed on 9 Sep 2019.
10. de Heer EW, Palacios JE, Adèr HJ, van Marwijk HWJ, Tylee A, van der Feltz-Cornelis CM. Chest pain, depression and anxiety in coronary heart disease: Consequence or cause? A prospective clinical study in primary care. J Psychosom Res. 2020;129:109891. https://doi.org/10.1016/j.jpsychores.2019.109891.
11. Dogan MV, Grumbach IM, Michaelson JJ, Philibert RA. Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study. PLoS ONE. 2018;13(1):e0190549. https://doi.org/10.1371/journal.pone.0190549.
12. Eom J, Kim S, Zhang B. AptaCDSS-E: A classifier ensemble-based clinical decision support system for cardiovascular disease level prediction. Expert Syst Appl. 2008;34(4):2465–79. https://doi.org/10.1016/j.eswa.2007.04.015.
13. Famili A, Shen W-M, Weber R, Simoudis E. Data preprocessing and intelligent data analysis. Intelligent Data Analysis. 1997;1(1):3–23. https://doi.org/10.3233/ida-1997-1102.
14. Gordon T, Castelli WP, Hjortland MC, Kannel WB, Dawber TR. High density lipoprotein as a protective factor against coronary heart disease. The American Journal of Medicine. 1977;62(5):707–14. https://doi.org/10.1016/0002-9343(77)90874-9.

15. Goyal M, Goyal R, Venkatappa P, Reddy BL. Activation functions deep learning: Algorithms and applications 2019;1–30 https://doi.org/10.1007/978-3-030-31760-7_1
16. Grover S, Bhartia S, Akshama Y, A., & K.R., S. Predicting severity of Parkinson's disease using deep learning. Procedia Computer Science. 2018;132:1788–94. https://doi.org/10.1016/j.procs.2018.05.154.
17. Haasenritter J, Stanze D, Widera G, Wilimzig C, Abu Hani M, Sönnichsen AC, Donner-Banzhoff N. Does the patient with chest pain have a coronary heart disease? Diagnostic value of single symptoms and signs – a meta-analysis. Croatian Medical Journal. 2012;53(5):432–41. https://doi.org/10.3325/cmj.2012.53.432.
18. Hung AJ, Chen J, Gill IS. Automated performance metrics and machine learning algorithms to measure surgeon performance and anticipate clinical outcomes in robotic surgery. JAMA Surgery. 2018;153(8):770. https://doi.org/10.1001/jamasurg.2018.1512.
19. Jagadish HV, Gehrke J, Labrinidis A, Papakonstantinou Y, Patel JM, Ramakrishnan R, Shahabi C. Big data and its technical challenges. Commun ACM. 2014;57(7):86–94. https://doi.org/10.1145/2611567.
20. Jindal U, Gupta S, Jain V, Paprzycki M. Offline Handwritten Gurumukhi Character Recognition System Using Deep Learning. In Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals. Springer: Singapore; 2020 p. 121-133.
21. Kantardzic M. Data mining: concepts, models, methods, and algorithms. John Wiley & Sons; 2011 Aug 16.
22. Kuutti S, Bowden R, Jin Y, Barber P, Fallah S. A survey of deep learning applications to autonomous vehicle control. arXiv preprint arXiv:1912.10773. 2019 Dec 23.
23. Lai Z, Deng H. Medical image classification based on deep features extracted by deep model and statistic feature fusion with multilayer perceptron□. Computational Intelligence and Neuroscience. 2018;2018:1–13. https://doi.org/10.1155/2018/2061516.
24. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–44. https://doi.org/10.1038/nature14539.
25. Lin W-C, Tsai C-F, Hu Y-H, Jhang J-S. Clustering-based undersampling in class-imbalanced data. Inf Sci. 2017;409–410:17–26. https://doi.org/10.1016/j.ins.2017.05.008.
26. Liu T, Ding X, Chen Y, Chen H, Guo M. Predicting movie box-office revenues by exploiting large-scale social media content. Multimedia Tools and Applications. 2014;75(3):1509–28. https://doi.org/10.1007/s11042-014-2270-1.
27. Liu W, Wang Z, Liu X, Zeng N, Liu Y, Alsaadi FE. A survey of deep neural network architectures and their applications. Neurocomputing. 2017;234:11–26. https://doi.org/10.1016/j.neucom.2016.12.038.
28. Ma X, Yu H, Wang Y, Wang Y. Large-scale transportation network congestion evolution prediction using deep learning theory. PLoS ONE. 2015;10(3):e0119044. https://doi.org/10.1371/journal.pone.0119044.
29. Martin Cichy R, Khosla A, Pantazis D, Oliva A. Dynamics of scene representations in the human brain revealed by magnetoencephalography and deep neural networks. NeuroImage. 2017;153:346–58. https://doi.org/10.1016/j.neuroimage.2016.03.063.
30. Masih N, Ahuja S. Prediction of Heart Diseases Using Data Mining Techniques. International Journal of Big Data and Analytics in Healthcare. 2018;3(2):1–9. https://doi.org/10.4018/ijbdah.2018070101.
31. Miikkulainen R, Liang J, Meyerson E, Rawal A, Fink D, Francon O, Raju B et al. Evolving deep neural networks. Artificial Intelligence in the Age of Neural Networks and Brain Computing, 2019;293-312. Academic Press.
32. Wlodarczak P, Soar J, Ally M. Multimedia data mining using deep learning. In 2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC), IEEE. p. 190-196 7 Oct 2015
33. Nahar J, Imam T, Tickle KS, Chen Y-PP. Association rule mining to detect factors which contribute to heart disease in males and females. Expert Syst Appl. 2013;40(4):1086–93. https://doi.org/10.1016/j.eswa.2012.08.028.
34. Nazarzadeh M, Pinho-Gomes A-C, Smith Byrne K, Canoy D, Raimondi F, Ayala Solares JR, Rahimi K. Systolic blood pressure and risk of valvular heart disease. JAMA Cardiology. 2019;4(8):788. https://doi.org/10.1001/jamacardio.2019.2202.
35. Nilashi M, Ibrahim O, Ahani A. Accuracy improvement for predicting Parkinson's disease progression. Scientific Reports. 2016;6(1):1–18. https://doi.org/10.1038/srep34181.
36. Niu J, An G, Gu Z, Li P, Liu Q, Bai R, Sun J, Du Q. Analysis of sensitivity and specificity: precise recognition of neutrophils during regeneration of contused skeletal muscle in rats. Forensic Sci Res. 2020. https://doi.org/10.1080/20961790.2020.1713432.
37. Palaniappan S, Awang R. Intelligent heart disease prediction system using data mining techniques. IEEE/ACS International Conference on Computer Systems and Applications. 2008;2008:108–15. https://doi.org/10.1109/aiccsa.2008.4493524.
38. Pandey A, Patel KV, Lavie CJ. Obesity, central adiposity, and fitness: Understanding the obesity paradox in the context of other cardiometabolic parameters. Mayo Clin Proc. 2018;93(6):676–8. https://doi.org/10.1016/j.mayocp.2018.04.015.
39. Pandi A, Koch M, Voyvodic PL, Soudier P, Bonnet J, Kushwaha M, Faulon J-L. Metabolic perceptrons for neural computing in biological systems. Nature Communications. 2019;10(1):1–13. https://doi.org/10.1038/s41467-019-11889-0.
40. Paynter NP, Balasubramanian R, Giulianini F, Wang DD, Tinker LF, Gopal S, Rexrode KM. Metabolic predictors of incident coronary heart disease in women. Circulation. 2018;137(8):841–53. https://doi.org/10.1161/circulationaha.117.029468.
41. Pencina MJ, Navar AM, Wojdyla D, Sanchez RJ, Khan I, Elassal J, Sniderman AD. Quantifying importance of major risk factors for coronary heart disease. Circulation. 2019;139(13):1603–11. https://doi.org/10.1161/circulationaha.117.031855.
42. Poplin R, Varadarajan AV, Blumer K, Liu Y, McConnell MV, Corrado GS, Webster DR. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. Nature Biomedical Engineering. 2018;2(3):158–64. https://doi.org/10.1038/s41551-018-0195-0.
43. Prescott E, Hippe M, Schnohr P, Hein HO, Vestbo J. Smoking and risk of myocardial infarction in women and men: Longitudinal population study. BMJ. 1998;316(7137):1043–7. https://doi.org/10.1136/bmj.316.7137.1043.
44. Prescott E, Hippe M, Schnohr P, Hein HO, Vestbo J. Smoking and risk of myocardial infarction in women and men: Longitudinal population study. 1998. https://doi.org/10.1136/bmj.316.7137.1043
45. Schmidhuber J. Deep learning in neural networks: An overview. Neural Networks. 2015;61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003.
46. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Hassabis D. Improved protein structure prediction using potentials from deep learning. Nature. 2020;577(7792):706–10. https://doi.org/10.1038/s41586-019-1923-7.
47. Singh MP, Chaturvedi S, Shudhalwar DD. Multilayer neural network technique for parsing the natural language sentences. International Journal of Artificial Intelligence and Machine Learning. 2019;9(2):22–38. https://doi.org/10.4018/ijaiml.2019070102.
48. Singh R, Srivastava S. Stock prediction using deep learning. Multimedia Tools and Applications. 2016;76(18):18569–84. https://doi.org/10.1007/s11042-016-4159-7.
49. Son C-S, Kim Y-N, Kim H-S, Park H-S, Kim M-S. Decision-making model for early diagnosis of congestive heart failure using rough set and decision tree approaches. J Biomed Inform. 2012;45(5):999–1008. https://doi.org/10.1016/j.jbi.2012.04.013.

50. Tajmiri S, Azimi E, Hosseini MR, Azimi Y. Evolving multilayer perceptron, and factorial design for modelling and optimization of dye decomposition by bio-synthetized nano CdS-diatomite composite. Environ Res. 2020;182:108997. https://doi.org/10.1016/j.envres.2019.108997.
51. Tang J, Deng C, Huang G-B. Extreme learning machine for multilayer perceptron. IEEE Transactions on Neural Networks and Learning Systems. 2016;27(4):809–21. https://doi.org/10.1109/tnnls.2015.2424995.
52. Tschandl P, Codella N, Akay BN, Argenziano G, Braun RP, Cabo H, Kittler H. Comparison of the accuracy of human readers versus machine-learning algorithms for pigmented skin lesion classification: An open, web-based, international, diagnostic study. Lancet Oncol. 2019;20(7):938–47. https://doi.org/10.1016/s1470-2045(19)30333-x.
53. Wang J, Ma Y, Zhang L, Gao RX, Wu D. Deep learning for smart manufacturing: Methods and applications. Journal of Manufacturing Systems. 2018;48:144–56. https://doi.org/10.1016/j.jmsy.2018.01.003.
54. Wang S-H, Zhang Y, Li Y-J, Jia W-J, Liu F-Y, Yang M-M, Zhang Y-D. Single slice based detection for Alzheimer's disease via wavelet entropy and multilayer perceptron trained by biogeography-based optimization. Multimedia Tools and Applications. 2016;77(9):10393–417. https://doi.org/10.1007/s11042-016-4222-4.
55. Whitaker K. Earlier diagnosis: The importance of cancer symptoms. Lancet Oncol. 2020;21(1):6–8. https://doi.org/10.1016/s1470-2045(19)30658-8.
56. Whitaker K. Earlier diagnosis: the importance of cancer symptoms. The Lancet Oncology. 2020;21(1):6-8.
57. Wlodarczak P, Soar J, Ally M. Multimedia data mining using deep learning. Fifth International Conference on Digital Information Processing and Communications (ICDIPC). 2015;2015:190–6. https://doi.org/10.1109/icdipc.2015.7323027.
58. Wlodarczak P, Jeffrey S, Mustafa A. Multimedia data mining using deep learning. In 2015 Fifth International Conference on Digital Information Processing and Communications (ICDIPC), IEEE: 2015;190-196.
59. Young T, Hazarika D, Poria S, Cambria E. Recent trends in deep learning based natural language processing [review article]. IEEE Comput Intell Mag. 2018;13(3):55–75. https://doi.org/10.1109/mci.2018.2840738.
60. Wolterink JM, Tim L, Max AV, Ivana I. Dilated convolutional neural networks for cardiovascular MR segmentation in congenital heart disease. In Reconstruction, segmentation, and analysis of medical images. Springer: Cham; 2016. p. 95-102.
61. Zoughi T, Homayounpour MM, Deypir M. Adaptive windows multiple deep residual networks for speech recognition. Expert Syst Appl. 2020;139:112840. https://doi.org/10.1016/j.eswa.2019.112840.
62. Zyriax B-C, Vettorazzi E, Hamuda A, Windler E. Interaction of smoking and dietary habits modifying the risk of coronary heart disease in women: Results from a case–control study. Eur J Clin Nutr. 2018;72(12):1673–81. https://doi.org/10.1038/s41430-018-0099-9.