CrossMark

SYSTEMS-LEVEL QUALITY IMPROVEMENT

# Role of Soft Computing Approaches in HealthCare Domain: A Mini Review

**Shalini Gambhir**[1] · **Sanjay Kumar Malik**[1] · **Yugal Kumar**[2]

**Abstract** In the present era, soft computing approaches play a vital role in solving the different kinds of problems and provide promising solutions. Due to popularity of soft computing approaches, these approaches have also been applied in healthcare data for effectively diagnosing the diseases and obtaining better results in comparison to traditional approaches. Soft computing approaches have the ability to adapt itself according to problem domain. Another aspect is a good balance between exploration and exploitation processes. These aspects make soft computing approaches more powerful, reliable and efficient. The above mentioned characteristics make the soft computing approaches more suitable and competent for health care data. The first objective of this review paper is to identify the various soft computing approaches which are used for diagnosing and predicting the diseases. Second objective is to identify various diseases for which these approaches are applied. Third objective is to categories the soft computing approaches for clinical support system. In literature, it is found that large number of soft computing approaches have been applied for effectively diagnosing and predicting the diseases from healthcare data. Some of these are particle swarm optimization, genetic algorithm, artificial neural network, support vector machine etc. A detailed discussion on these approaches are presented in literature section. This work summarizes various soft computing approaches used in healthcare domain in last one decade. These approaches are categorized in five different categories based on the methodology, these are classification model based system, expert system, fuzzy and neuro fuzzy system, rule based system and case based system. Lot of techniques are discussed in above mentioned categories and all discussed techniques are summarized in the form of tables also. This work also focuses on accuracy rate of soft computing technique and tabular information is provided for each category including author details, technique, disease and utility/accuracy.

**Keywords** Artificial intelligence · Healthcare database · Medical data mining · Soft computing · Medical diagnosis

✉ Yugal Kumar
yugalkumar.14@gmail.com

Shalini Gambhir
shalinigambhir.21@gmail.com

Sanjay Kumar Malik
skmalik9876@gmail.com

[1] Department of Computer Science and Engineering, SRM University, Delhi NCR, Sonipat, Haryana, India

[2] Department of Information Technology, KIET Group of Institution, Ghaziabad, India

## Introduction

Large quantity of data is generated in health care domain and rate of storing healthcare information in clinical databases is growing with a fast pace. Hence, it becomes important to find the relevant information from the explosive amount of data for effective treatment of patients and diseases diagnosis and prediction. Presently, research community give more attention towards healthcare domain data and tries to establish the relationship between health care data and machinery learning algorithms and mine the significant information for better treatment of diseases. Large number of machine learning and medical data mining model have been presented in literature. Nowadays, application of soft computing techniques are also increased gradually in diverse areas of medical services such

Springer

as for helping the medical expert, owing to effectiveness and continuous improvement of classification and detection system, upgrading diagnostic rate and many more.

## Data mining and soft computing

A non trivial process of extracting potentially profitable, novel, valid and ultimately understandable patterns from huge collection of data is termed as knowledge discovery [1]. According to Cios, et al. [2] and Cios & Moore [3] knowledge discovery process is a six step mode, which involves multiple two-way iterations and interactions, as it may impact other steps in the series. The six steps of knowledge discovery process can be shown in Fig. 1 and consist of feedback loops which shows the continuous improvement in the model. Whereas, the block diagram for the knowledge discovery process is shown in Fig. 2. The objective of knowledge discovery process is to define a set of steps such as plan, work through and reduce the cost of project, describing procedures to be undertaken as part of each step. But, due to rapid increasing of numbers of existing databases and size of information, it is tough task to extract the the information manually from these databases [5]. Data mining is one of the most important processes for finding valuable information and patterns from databases and further this extracted information can be used in

decision making process. It is iterative in nature and follows an interactive methodology, including many states such as visualization of information, various database techniques, recognition of patterns, neural networks, artificial intelligence, machine learning, statistical techniques, etc. Recently, various soft computing techniques have also been encapsulated along with data mining to handle most vital, challenging and motivating areas of medical research. As we are aware, the last decade has witnessed major advancements in the methods of the diagnosis. Soft computing techniques involve multiple methodologies working synergistically to provide significant information processing capacities for managing realistic but ambiguous situations. The prime challenges in traditional approaches are mainly precision, certainty as well as hardness. But, in soft computing, the principle notion is to devise methods of calculation leading to evolution of low cost solution. It is found that the adoption of soft computing for data mining task improve the accuracy rate as well as reduce the time constraint. Earlier, data processors were involved in making a knowledge based decision support systems in the clinical area, collecting data from medical experts and transferring in the strain of computer algorithms manually, thus ultimately leading to knowledge creation. Apparently, it became a time consuming process and also include the manual processes of medical expert opinion. To overcome above mentioned

**Fig. 1** Six step knowledge discovery process
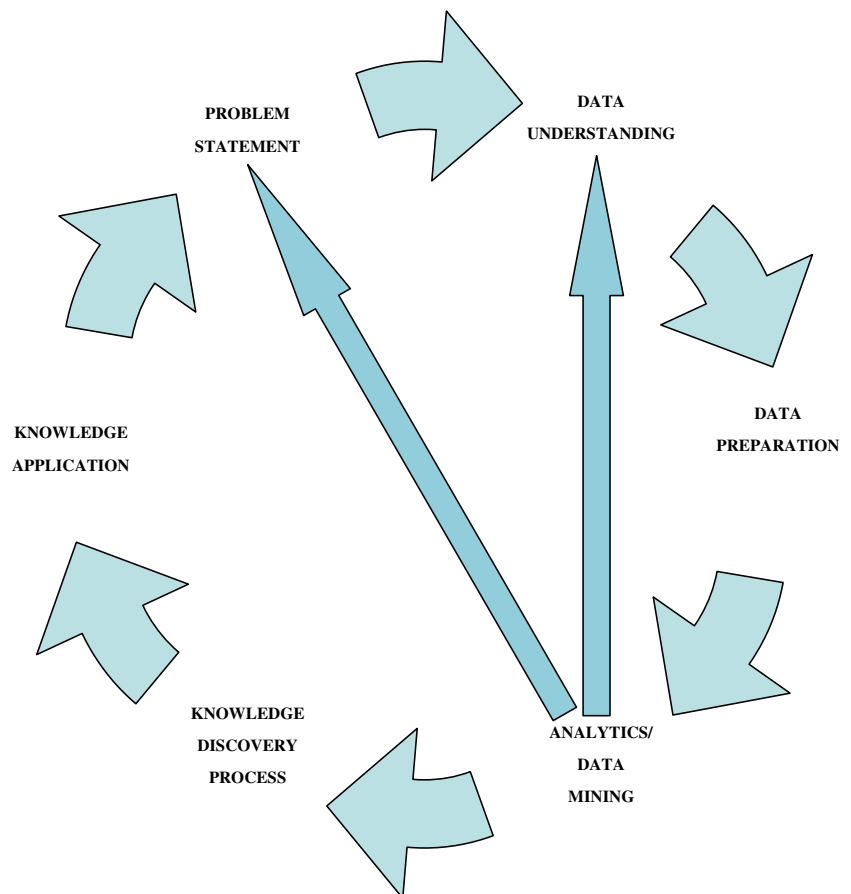
problem is to applicability of soft computing technologies in medical expert opinion such as genetic algorithms, neural nets, fuzzy logics, etc. [6–9].

Further, in developing countries like India, medical challenges are increasing day by day with more crunches on resources due to ever increasing population. Shortage of medical specialists further adds up to the problem of high mortality and various diseases. Sometimes, it is also seen that patients ignore the symptoms of diseases in earlier stage and approach for medical facilities either when it's too late or it become extremely critical. One of the solution for handling such types of problems is to use intelligent soft computing prediction model over medical data and build medical support organizations which can mimic the doctor's reasoning and conclude the symptoms without consulting of specialists. The integration of the computer technology with medical field is one of the most dynamic research area.

## Method for literature review

### Research question (RQ)

The primary aim of this review is to find the answer of following research questions (RQ). These are

> RQ1: What are the various soft computing techniques available for diseases diagnosis and prediction?
> RQ2: What is the overall accuracy achieved from these techniques?
> RQ3: What are the existing research gaps in context of soft computing techniques and disease diagnosis and prediction?
> RQ4: Which technique can improve the diagnosis or prediction rate for physician's decision-making process?

### Source of information

Due to wide scope of review articles, it is recommended that various research databases can be explored for existing work. So in this study, six databases are searched to find the existing
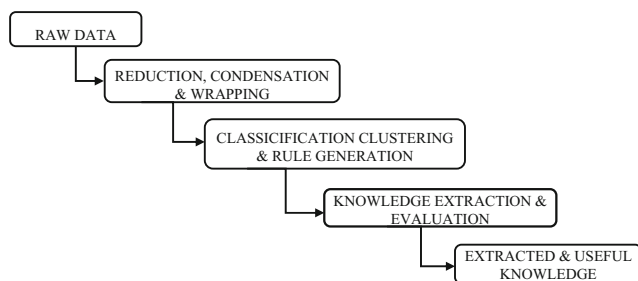


**Fig. 2** Block diagram for knowledge discovery

related work in the domain of soft computing and medical data mining. These databases are

- Google scholar (scholar.google.com)
- IEEE eXplore (www.ieeexplore.ieee.org)
- Springer (www.springerlink.com)
- Science Direct (www.sciencedirect.com)
- ACM digital library (dl.acm.org)
- Mendeley

### Inclusion and search criteria

The aim of this study is to outline the relationship between various soft computing techniques and healthcare domain particularly, the applicability of soft computing techniques for diseases diagnosis and prediction. Hence, the studies based on diseases diagnosis and predictions using soft computing techniques are included in this review if the following criteria is met.

- Related to soft computing techniques
- Includes the data on diseases diagnosis and prediction
- Computerized implementation and systems to detect diseases
- Includes single and multiple occurrence of diseases
- Work published in between 2008 to 2016
- Published in SCI listed journals

Initial search considers all relevant existing work in the area of soft computing techniques and diseases diagnosis and prediction. The SLR of this study is mentioned in Fig. 3. The keywords used for this work are (((Soft Computing Techniques) < OR > (Machine Learning) < OR > (Evolutionary methods) < OR > (Swarm Based Techniques)) < AND> ((Diseases diagnosis) < OR> (Diseases Prediction) < OR > (Diseases detection)) < AND > ((Classification Algorithm) < OR> (Classification Model) < OR> (Expert System) < OR> (Rule Based System) < OR> (Fuzzy System) < OR> (Neuro Fuzzy System) < OR> (Cases based Reasoning))).

To find all existing literature work, the above search string is applied on full text rather than title or abstract. The range of the literature work is set to January 2008 up till February 2016. In this study, various research databases are searched for the existing literature survey. The different research databases having different searching standards. Hence, the searching strategy is also designed accordingly. Some of databases do not have the "And" option. For those, search string is reassessed and the search can perform using following keywords.

1) ((Soft Computing Techniques) < OR > (Machine Learning) < OR > (Evolutionary methods))

**Fig. 3** Methodology and evaluation of review on soft computing techniques and health care data
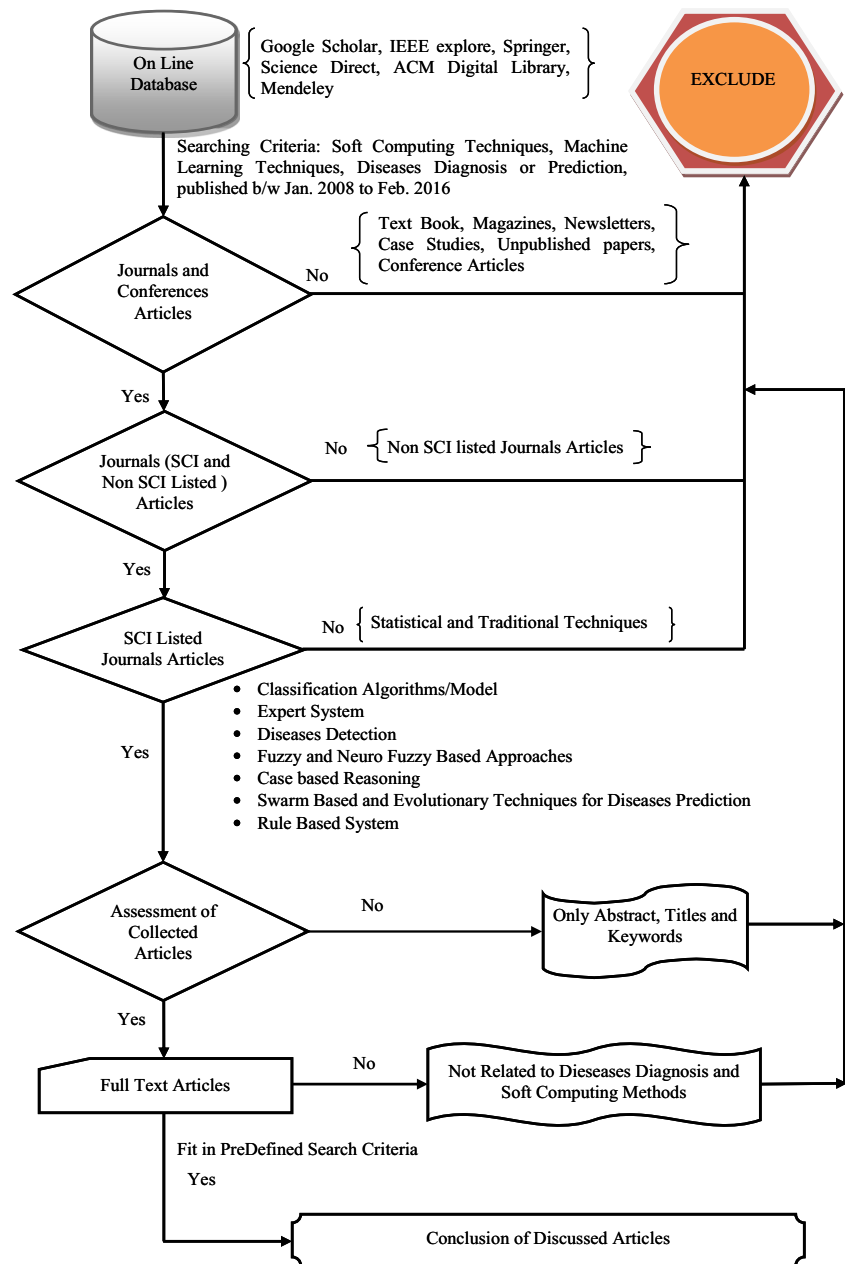


2) ((Diseases diagnosis) < OR > (Diseases Prediction) < OR > (Diseases detection))

3) ((Classification Algorithm) < OR > (Classification Model) < OR > (Expert System) < OR > (Rule Based System) < OR > (Fuzzy System) < OR > (Neuro Fuzzy System) < OR > (Cases based Reasoning))

**Exclusion criteria**

The exclusion criteria is also considered after search process. The exclusion criteria includes content from books, magazines, newsletters and educational courses; case studies;

articles published in national and international conferences, symposium and workshops and articles published in non SCI listed journals.

**Extraction of articles**

Initially, 2658 research articles are collected from all sources mentioned in subsection 2.2. Huge amount of research articles are collected due the generic terms "soft computing techniques" and "diseases diagnosis and prediction". One of the reason is that the research databases also search the articles using computing, techniques, methods and detection keywords. In the next step, articles irrelevant to our search criteria

i.e. soft computing techniques and diseases diagnosis, are removed. As a result of this, only 1784 articles are left. Further, In this study, the articles published in SCI listed journals are considered. But, the search string is not capable to Differiante between SCI listed journals and non SCI listed journals. So, this can be done in manually and all of articles published in non SCI journals, books and magazines are removed from the list, resulting in the exclusion of 1560 articles. Rest of article is assessed in terms of title, abstract, full text and keywords. As a result of this, only 118 articles are considered for review. Further, it is also noticed that 26 article are not fitted well in our predefined search criteria and these are articles are excluded from the list. Finally, 95 articles are undergone for the review study. All these articles have been published in SCI listed journals and Table 1 shows the details of journals, publishers and number of articles being published in journal.

Further, for scrutiny of collected articles, a team of five researchers is being formed. This team is responsible for the selection of the research articles using our predefined search criteria. Initially, three researchers select the articles and the selected articles are further cross checked through fourth and fifth members of team. The conflict situation can handle using collectively decision of all team members. This process is repeated at each step of study selection. It is also found that twenty three different diseases are diagnosed and predicted using soft computing techniques. Out of 95 research articles, 36 articles have been published with new technology, whereas, 47different feature selection/ parameters optimization and rule generation/ extraction methods have also been reported in finalized articles. It is also observed that the SVM techniques are widely employed for disease diagnosis and prediction. This technique can be applied in fifteen different research articles whereas ANN techniques can be used in eleven

**Table 1** Categorization of research article with journals

| Sr. No. | Journal Name | Publisher | No. of Papers |
|---|---|---|---|
| 1 | Knowledge Based Systems | Elsevier | 5 |
| 2 | Expert Systems with Applications | Elsevier | 34 |
| 3 | Egyptian Informatics Journal | Elsevier | 1 |
| 4 | Engineering Applications of Artificial Intelligence | Elsevier | 1 |
| 5 | Neural Networks | Springer | 2 |
| 6 | Neurocomputing | Elsevier | 2 |
| 7 | Artificial Intelligence in Medicine | Elsevier | 5 |
| 8 | Journal of Biomedical Informatics | Elsevier | 2 |
| 9 | Journal of Medical Systems | Springer | 8 |
| 10 | Computer Methods and Programs in Biomedicine | Elsevier | 3 |
| 11 | IEEE Transactions on Information Technology in Biomedicine | IEEE | 2 |
| 12 | IEEE Transactions on Systems Man and Cybernatics | IEEE | 2 |
| 13 | Computers in Biology and Medicine | Elsevier | 2 |
| 14 | Applied Soft Computing | Elsevier | 5 |
| 15 | Medical Engineering & Physics | Elsevier | 1 |
| 16 | International Journal of Approximate Reasoning | Elsevier | 1 |
| 17 | Fuzzy Sets and Systems | Elsevier | 1 |
| 18 | Soft Computing | Springer | 3 |
| 19 | Neuropsychologia | Elsevier | 2 |
| 20 | European Journal of Operational Research | Elsevier | 1 |
| 21 | International Journal of Medical Informatics | Elsevier | 1 |
| 22 | Archives of Gerontology and Geriatrics | Elsevier | 1 |
| 23 | Cluster Computing | Springer | 1 |
| 24 | Cybernetics and Systems | Taylor Francis | 1 |
| 25 | Technology and Health Care | IOS Press | 1 |
| 26 | IEEE TRANSACTIONS ON MEDICAL IMAGING | IEEE | 1 |
| 27 | Neural Computing and its Applications | Springer | 2 |
| 28 | International Journal of Computer Assisted Radiology and Surgery | Springer | 1 |
| 29 | Expert Systems | Willey | 1 |
| 30 | Decision Support System | Elsevier | 1 |
| 31 | Computers in Industry | Elsevier | 1 |

different articles. PSO is most significant technique which can be applied for parameter optimization of SVM and ANN, whereas, PCA is most reliable technique for features reduction.

## Data extraction strategy

All the finalized research articles are categorized in five different techniques and also explored thoroughly to find some common key properties for comparison. The data extracted from these articles are presented in Tables 1, 2, 3, 4, 5 and 6, and the five different architecture based on the information contained in these articles are shown in Figs. 4, 5, 6, 7 and 8. Each article studies through and evaluates to find the following.

- Technique/Methodology Adopted
- Type of Disease
- Utility/ Accuracy
- Authors

## Literature review

This section describes the literature survey related to usage of soft computing techniques in medical data analysis. This section includes various techniques related to classification, clustering, regression and association for disease prediction and diagnosis. For this study, substantial research literature related to mining on medical data using soft computing is collected. All collected literature divides into five subsections on the basis of information and data. Each subsection summarizes the latest research or progress in the era of soft computing techniques and its applicability in healthcare domain. Figure 4 shows the various categories of clinical support system used in this study.

### Classification method based systems

Classification is a simple data mining method to divide the sample items into target categories with a goal for ensuring that the each items belongs to its category i.e. close to accuracy. For example, classification can be used to determine the condition of patient is either better or worst on a particular day based on disease symptoms. The classification method can be broadly categorized into a binary and multilevel classification. The most elementary model for classification problem, when only two potential values for the target attribute exists, known as binary classification. For instance, high credit rating or low credit rating. While, in multilevel approach, multiple values are presented in terms of target class such as low, medium, high, or unknown credit rating. In classification, information

set is divided into training and test sets, the training set is used to make the model, whereas, the test set is used to square off the accuracy of the model and also to validate training model. Classification model is one of the commonly used models applicable in medical data mining. Figure 5 shows the process of classification method based systems. Table 2 summarizes the information on medical data mining techniques based on classification methods.

For automatic detection of normal and Coronary Artery Disease conditions, Giri et al. [10], have used heart signals rather than cardic signals to design a methodology [10]. The heart signals are associated with several frequency sub bands using Discrete Wavelet Transform (DWT). Hence, to reterive the heart signal from DWT and also to reduce the dimension of data, three statistical methods are applied on the set of DWT coefficients. Further, the normalized information is passed through a set of four classifier such as Support Vector Machine (SVM), Probabilistic Neural Network (PNN), Gaussian Mixture Model (GMM), and K-Nearest Neighbor (KNN) for decision making. The results indicate that the GMM model coupled with ICA provides higher accuracy rate. Further, to improve the classification performance of the Coronary Artery Disease, Babaoglu et al., have developed a BPSO, GA and SVM based classification model based on stress testing data [11]. In the proposed model, BPSO and GA techniques are applied for obtaining the relevant set of features to predict Coronary Artery Disease. Whereas, SVM is adopted as classifier system to classified patients with CAD. For enhancing the diagnostic rate, 10 cross fold technique is incorporated in SVM classifier. Results indicate that the proposed classification model achieves higher accuracy rate and having less complexity. To investigate the characteristics of diabetic patients and also to predict type-2 diabetic effectively, Patil et al., have developed a hybrid prediction model for diabetic disease [12]. In the proposed prediction model k-means clustering algorithm is applied to validate the class labels of diabetic dataset and classical C4.5 method is selected to build the final classifier using k-cross fold validation method. From results, it is stated that the proposed system attains higher sensitivity and specificity rate than the other methods being compared. An intelligent hepatitis diagnostic system for caring and treatment of hepatitis patients is reported in [13]. The working of the proposed diagnostic system is divided into two parts- features extraction and reduction, and classification of hepatitis disease. In the features extraction and reduction part, firstly features are extracted from hepatitis database and further, the more relevant features are obtained by using PCA method. Whereas, in classification part, LSSVM classifier is employed on reduce set of features to obtain final results. It is seen that the PCA-LSSVM based diagnostic system not only improve the accuracy rate of hepatitis disease but also enhance the sensitivity and specificity rate. Another application of SVM technique is also reported for effective diagnosis of

**Table 2** Summary of medical data mining techniques using Classification Approach

| Author | Techniques | Disease | Utility/Accuracy |
|---|---|---|---|
| Giri et al. [10] | SVM, Probabilistic Neural Network, GMM, and k-NN | Coronary Artery Disease | Highest accuracy |
| Babaoglu et al. [11] | BPSO, GA and SVM | Coronary Artery Disease | Clssification Model |
| Patil et al. [12] | K-means and C4.5 algorithm | Type-2 Diabetic Disease | Hybrid Prediction Model (HPM) |
| Calisir et al. [13] | PCA and LSSVM | Hepatitis Disease | Expert Hepatitis Diagnosis System |
| Sweilam et al. [14] | QPSO and LSSVM | Cancer Disease | Classification Tool |
| Abdi et al. [15] | PSO and SVM | Erythemato-Squamous Diseases | 98.91 % Classification |
| Bhardwaj A. [16] | GA and ANN | Breast Cancer | Classification accuracy of 98.24 %, 99.63 % and 100 % by 50–50, 40–60, 30–70 testing-training partition |
| Fei et al. [17] | PSO and SVM | Arrhythmia Disease | PSO-SVM Classification Model |
| Chen et al. [18] | LFDA and SVM | Hepatitis Disease | Hybrid Classification Model |
| Nasiri et al. [19] | LVQ , ALVQ-RM and ANN | Breast cancer and Thyroid diseases | Automatic disease diagnosis systems |
| Lopez et al. [20] | PCA, LDA, FDR, NN and SVM | Alzheimer's Disease | Computer Aided Diagnosis System |
| Park et al. [21] | GA, C5.0 and CART | Heart, Diabetes, Hepatitis and Breast Cancer Diseases | Cost Sensitive Based Classification System |
| Latifoglu et al. [22] | PCA, k-NN and AIRS | Atherosclerosis Disease | 100 % Classification Accuracy |
| Zheng et al. [23] | K-means and Support Vector Machine | Breast Cancer Disease | 97.38 % Accuracy |
| Yeh et al. [24] | Decision tree, Bayesian classifier and Back propagation neural network | Cerebrovascular Disease | Classification model with sensitivity of 99.48 % and specificity of 99.59 %. |
| Eom et al. [25] | AptaCDSS-E system based on SVM, NN, DT and BN | Cardiovascular Disease (CVD) | Diagnositic accuracy >94 % |
| Ubeyli et al. [26]] | K-means clustering algorithm | Erythemato-Squamous Disease | Higher Classification Accuracy |
| Ozcift et al. [27] | Rotation Forest (RF) Ensemble Classifier | Parkinson's, Diabetes and Heart Diseases | Computer-Aided Diagnosis (CADx) Systems with improve accuracy |
| Jerez et al. [28] | Multi-layer perceptron (MLP), Self organisation maps (SOM) and K-nearest neighbour (KNN) | Breast Cancer Disease | Significant enhancement in prognosis accuracy |
| Das et al. [29] | Bagging, Boosting and Random subspace | Valvular Heart Disease | Ensemble classification methods with improve detection rate |
| Sartakhti et al. [30] | Support Vector Machine (SVM) and Simulated Annealing (SA). | Hepatitis Disease | Classification accuracy > 96 % |
| Nahar et al. [31] | NaiveBayes, IBK and SMO | Heart Disease | Promised results for heart disease diagnosis |
| Morra et al. [32] | AdaBoost and SVM | Alzheimer's Disease | Ada-SVM perform better than AdaBoost and FreeSurfer |
| Er et al. [33] | Learning Vector Quantization, and Generalized Regression Neural Networks | Chest Diseases | Better diagnostic rate for chest disease |
| Er et al. [34] | MLNN with one hidden layer, MLNN with two hidden layers and General Regression Neural Network (GRNN) | Tuberculosis Disease | Better diagnostic rate using GRNN |
| Kampouraki et al. [35] | Support Vector Machine | Heartbeat | Classification Model |
| Elveren et al. [36] | MLNN and GA | Tuberculosis Disease | Enhanced tuberculosis disease classification rate |
| Cho et al. [37] | LRBF, VRIFA and SVM | Breast Cancer, Diabetes, and Heart disease | Automatic diagnosis system |
| Kumar and Inbarani [101] | Rough Set, DWT | Cardic Disease | 99.8 % Classification Accuracy |
| Prasad et al. [103] | Rough set and Machine learning | Thyroid Disease | 99 % Accuracy rate. |

**Table 3**    Summary of medical data mining techniques using Expert Systems

| Author | Techniques | Disease | Utility/Acuuracy |
|--------|-----------|---------|------------------|
| Lee C.S. et al. [38] | Layer Fuzzy Ontology | Diabetes | Decision Support System and it is more effective in the Diabetes application |
| Sengur A. [39] | Linear Discriminant Analysis and Adaptive Neuro-Fuzzy Inference System | Heart valve Disorder | 95.9 % Sensitivity and 94 % Specificity rate |
| Sengur A. [40] | Principal Component Analysis, Artificial Immune System and Fuzzy k-NN | Heart valve Disorder | Obtained 95.9 % sensitivity and 96 % specificity rate |
| Pal D. et al. [42] | Fuzzification of clinical parameters and Defuzzification of fuzzy output to crisp value | Coronary Artery Disease (CAD) | 95.85 % Sensitivity and 83.33 % Specificity accuracy |
| Hariharan et al. [43] | PCA, LDA, LS-SVM, PNN and GRNN | Parkinson's Disease | Hybrid intelligent system with 100 % classification rate |
| Keles A. et al. [44] | Neuro Fuzzy | Thyroid diseases. | 95.33 % accuracy thyroid diseases. |
| Chen H.L. et al. [45] | K-Nearest Neighbor (FKNN) | Parkinson's Disease (PD) | 96.07 % classification accuracy |

cancer affected patients [14]. In this work, QPSO algorithm is applied to train the various parameters of SVM and an active set strategy is also introduced for obtaining good results.

To accurate diagnosis of erythemato-squamous disease, Abdi et al. [15], have developed a diagnosis model based on particle swarm optimization (PSO), support vector machines (SVMs) and association rules (ARs). In this model, ARs are used to to determine an optimal subset of features. Whereas, PSO algorithm is adopted to optimize the various parameters of SVM. Finally, the optimized SVM is applied to identify the patinets with erythemato-squamous disease or not. It is also seen that the proposed model have 98.91 % accuracy rate. For early detection and effective treatment of breast cancer patients, Bhardwaj et al. [16], have proposed a genetic algorithm inspired neural network model for diagnosing the breast cancer. In their work, authors optimize the value of weight and neural netwok structure by using genetic algorithm. The results of the proposed model are compared with BPNN and Koza model of neural network. It is observed that genetically optimized neural network provides better results. Fei has presented an effective classification model based on PSO and SVM techniques for effective treatment of arrhythmia disease [17]. In this model, PSO is integrated within SVM to optimize its parameters. The results illustrate that PSO-SVM model attains higher diagnostic rate in comparsion to ANN for the diagnosis of arrhythmia cordis. For the better diagnostic rate of hepatitis disease, Chen et al., have described a hybrid classification algorithm based on fisher discriminant analysis (LFDA) and supporting vector machine [18]. In their work, authors adopt the LFDA method as a feature selction tool, whereas SVM is applied to improve the diagnostic rate of hepatits disease. It is noticed that proposed hybrid classification algorithm provides higher accuracy rate i.e. 96.7 %. An automated disease classification system has been reported in [19]. In their work, author adopts LVQ based ANN for effective diagnosis and treatments of various dieseases. The reults report that LVQ-ANN gives better reults in comparsion to other established studies. A computer aided diagnosis system is developed for automatically detection of Alzheimer's disease (AD) [20]. In this work, neuroimages of brain are considered for the evaluation of Alzheimer's disease. It is also noted that PCA method is adopted for finding the relevant features from neuroimages and these features are further enhanced by applying LDA technique. Finally, NN and SVM methods are employed to detect the Alzheimer's disease affected patients. Results stated that the NN and SVM models provide significant improvement in accuracy rate as compared to classical voxels-as-features (VAF) reference approach. To overcome the shortcomings of classical CBR systems, Park et al., have presented a cost sensitive based system for disease diagnosis, called CSCS [21]. In proposed system, genetic algorithm is adopted to find the optimal classification cut off points. The results of CSCS are comparing to two other cost sensitive methods i.e. C5.0 and CART. It is found that the CSCS method gives overall better results than C5.0 and CART.

For the diagnosis of atherosclerosis, Latifoglu et al., have developed a medical diagnostic system based on PCA, k-NN and artificial immune recognition system (AIRS) [22]. In this work, carotid artery doppler signals are considered to diagnosis the atherosclerosis disease. Further, it is also noted that the proposed medical diagnostic system consists of four phases. In first phase, fast Fourier transformation (FFT) technique is applied to obtain the features of disease. In second phase, PCA techniques is applied to reduce the dimensions of data as well as to find more suitable and important features from features set. In third phase, a k-NN based weighted technique is adopted for preprocessing of data. Finally in fourth phase, AIRS classifier is applied to find the patients with atherosclerosis. The proposed system obtains almost 100 % accuracy rate.

**Table 4** Summary of medical data mining techniques using Fuzzy and Neuro- fuzzy Based Approach

| Author | Techniques | Disease | Utility/Accuracy |
|---|---|---|---|
| Ucar et al. [46] | Adaptive neuro fuzzy inference system and Rough sets | Tuberculosis Disease | 97 % Accuracy rate |
| Uguz [47] | Adaptive neuro-fuzzy inference system | Heart Valve diseases | 98.33 % Classification accuracy |
| Muthukaruppan et al. [48] | Particle swarm optimization (PSO) -based Fuzzy Expert System | Coronary artery disease (CAD) | 93.27 % Accuracy |
| Seera and Lim [49] | Fuzzy Min–Max neural network, Regression Tree and Random Forest | Breast Cancer, Diabetes and Liver Disorders | Hybrid Intelligent System (FMM-CART-RF) |
| Ubeyli et al. [50] | Adaptive Neuro-fuzzy Inference System (ANFIS) | Breast Cancer | Detecting the breast cancer |
| Papageorgiou et al. [52] | Cognitive Maps and Fuzzy Rule-Extraction methods. | Radiation therapy in Cancer Patients | Dynamic Decision System |
| Sanz et al. [53] | Fuzzy system | Cardiovascular Diseases | Best performance |
| Pal et al. [54] | Fuzzy rule-based system | Coronary artery disease | Screening System with 95.85 % sensitivity rate |
| Kannathal et al. [55] | Neuro fuzzy | Heart Diseases | 85–100 % accuracy |
| Khatibi and Montazer. [56] | Dempster–Shafer theory and fuzzy sets theory | Coronary artery disease (CAD) | Heart disease risk assessment tool |
| Uzoka et al. [57] | Analytic hierarchy process (AHP) and Fuzzy Inference | Malaria | Fuzzy–AHP Decision Support System |
| Long et al. [100] | Fuzzy C-Means, Rough Set, FireFly algorithm | Heart Diseases | Higher classification rate |

## Expert systems

An expert system is an application to help healthcare providers to make clinical decisions by analyzing the medical data. It can be understood as a computer program simulating judgement and tries to imitate human or an organization that has experience as well as expert knowledge in a particular subject area. The usage of artificial intelligence methods is increasing because of the effectiveness in classification and detection schemes assisting experts in the medical field. The architecture of expert system is explained in Fig. 6 and Table 3 provides the utility of different expert system based approaches.

It is well know fact that the classical systems are not handle imprecise and vague information effectively. The classical systems work with sharp boundaries. But, the fuzzy systems have capability to represent the imprecise and vague information. Hence, to keep in mind the application of fuzzy theory, Lee et al., have developed a fuzzy based expert system for diabetic disease [38]. In their work, authors have proposed a five layer fuzzy ontology based expert system for effective diagnosis of diabetic disease and provides remarkable results. Due to increased applicability of artificial methods in medical diagnosis, Sengur has developed a heart valve disorder detection system to determine the normal and abnormal heart valves [39]. In this study, doppler heart sounds are considered for the detection of normal and abnormal heart valves patients. The proposed system consists of three layers. First layer performs data pre processing tasks viz. filtering, normalization and denoising etc. In the second layer, features are extracted from doppler heart sounds using wavelet packet decomposition method and further, the dimension reduction is also done with the help of PCA technique. The third layer of the system is responsible to classify the patients with normal and abnormal heart valves. For the classification pupose, artificial immune system (AIS) and fuzzy k-NN techiques are implemented in third layer. Results report that the proposed detection system provides higher sensitivity and specificity rate. In continuation of their work, sengur has also developed another detection system for heart valve disorder using linear discriminant analysis (LDA) and adaptive neuro-fuzzy inference system (ANFIS) [40]. To exchange the information within medical devices, Arsene et al., have developed collaborative expert system for medical diagnosis [41]. The proposed expert system is a collabrative framework of the software agents and Bayesian networks are applied for repersentaion of the knowledge. This system consist of three key concept i.e. knowledge management, uncertainty reasoning and software components. Authors concluded their work through hybrid arcitecture of AI techniques, software agents and Bayesian network. In medical science, it is still tough task to identify earlier symptoms of diseases. To address the same, Pal et al., have developed a prescreening expert system for diagnosis of coronary artery disease (CAD) in earlier stage [42]. In their

**Table 5**  Summary of medical data mining techniques using Rule Based Approach

| Author | Techniques | Disease | Utility/Accuracy |
| --- | --- | --- | --- |
| Ell et al. [63] | Fuzzy Logic | Parkinson's Disease | Rule-based model with working memory and suboptimal decision strategies for PD affected patients |
| Astrom and Koker [64] | Parallel feed-forward neural network and Rule Based System | Parkinson's Disease | Higher classification accuracy |
| Kong et al. [65] | Belief rule-based inference method | Cardiac Chest Pain Disease | Clinical decision support system with belief rule based prototype |
| Lisboa et al. [67] | Rule extraction | Breast cancer | Neural network distinguish patients at the extremes of high and low risk. |
| Price et al. [69] | Rule-based learning | Parkinson's Disease | Rule-based category learning depends upon a network of strata structures and frontal cortical regions |
| Seto et al. [70] | Rule-based expert system | Heart Disease | Rule set increased self-care and improved the clinical management of pump failure |
| Wei.et al. [71] | Two stage feature selection | Orthopedic Disease | Accuracy obtained by using three sub-datasets are 68.82 %, 92.54 %, and 91.99 % |
| Sevastianov et al. [72] | Rule based and Evidential reasoning | Type 2 Diabetes Disease | Evidential reasoning based framework with improved classification rate |
| Nahar et al. [73] | Apriori, Predictive Apriori and Tertius rule generation methods | Heart disease | Association rule mining based heart disease prediction system |
| Jung et al. [74] | Rule based algorithm | Chronic Disease | Clinical decision support system |
| Toro et al. [75] | Rule Extraction using DDNA and SOEKS | Alzheimer's Disease | Experience-based clinical decision support system |
| Chen et al. [76] | Semantic Web Rule Language (SWRL) and Java Expert System Shell (JESS | Diabetes Mellitus | Diabetes Medication Recommendation system |
| Stoean and Stoean [77] | SVM and EA | Breast Cancer Disease | A two-step hybridized rule based model with high prediction accuracy |
| Sarkar et al. [78] | Decision tree based rule extraction and GA | Heart Disease | Rule based system with improved accuracy |
| Ang et al. [79] | EMA based rules extraction and AIS | Cancer and Diabetes Diseases | Performance of EMA based rules extraction model is more effective and significant |
| Mohamed et al. [80] | GA based rule construction method and ANN | Breast Cancer Disease | Adopt GA to extract the rules for artificial neural network technique and provides more efficient results |
| Kumar and Sahoo [81] | Decision tree for rule formation | Liver Disease | Rule based classification model for prediction of different types of liver disease. |

work, authors identify the risk factors responsible for CAD and further, these risk factors are derived in fuzzy rules with help of doctors and fuzzy inference system. Finally, prescreening expert system is presented for diagnosis of coronary artery disease. It is observed that the proposed system is capable to identify risk of CAD at an earlier stage. Further, it is noted that the proposed system achieves higher sensitivity rate up to 95.85 %. Hariharan et al., have presented an expert diagnostic system for the prediction of Parkinson's Disease [43]. In this work, PCA and LDA methods are used to extract features and the extracted features, SFS and SBS methods are used for dimension reduction and selected attributes are further classified with the help of LS-SVM, PNN and GRNN to find patients with PD. The proposed diagnostic system attains higher accuracy rate of 100 %.

For earlier detection and effective treatment of thyroid disease, Keles et al., have presented a thyroid disease expert system, called it ESTDD [44]. In ESTDD, neuro fuzzy classification method is adopted for detection of thyroid disease. It is found that ESTDD system gains 95.33 accuracy rate to diagnose thyroid disease. To improve the detection rate of Parkinson's disease (PD) and also making an effective and efficient diagnosis system for PD, Chen et al., have presented a diagnostic expert system for PD affected patients [45]. In this work, fuzzy k-nearest neighbor method is employed for Parkinson's disease diagnosis. To improve the detection rate, PCA technique is adopted to find the new features for Parkinson's disease. From results, it can be seen that the proposed diagnostic system obtains 96.07 % accuracy rate using 10 cross fold validation method.

### Fuzzy and neuro-fuzzy based system

Fuzzy or Neuro fuzzy based systems that are used in health check decision making provide more fleshed out medical exam data to be examined in a briefer time. In the last two decades, the use of fuzzy method of medical analysis is increasing as classification and detection systems have enhanced a great bargain to serve the medical experts in the assignments. The block diagram of fuzzy and neuro-fuzzy based system is presented in Fig. 7. Table 4 depicts the information of different fuzzy and neuro fuzzy based systems.

By adopting the application of fuzzy set theory, Ucar et al., have proposed a new hybrid machine learning method for identification of tuberculosis disease [46]. The proposed machine learning method is the combination of Adaptive neuro fuzzy inference system and Rough sets. The results indicate that the proposed method provides more vibal results in comprasion to other algorithms being compared.

To predict the heart valve disorder, Uguz, has developed an adaptive neuro fuzzy based system [47]. In the proposed system, three layers are presented for features extraction, features selection and classification. Features extraction is done by DWT, whereas, Shannon entropy algorithm is adopted for features selection. Finally, all selected features are classified using ANFIS classifier. The proposed system obtains 98.3 % classification accuracy. For the effective diagnosis of coronary artery disease (CAD), Muthukaruppan et al., have presented particle swarm optimization (PSO)-based fuzzy expert system [48]. In the proposed system, initially decision tree is implemented to find the best features for better prediction of coronary artery disease. Further, these features are coverted into fuzzy if then rules and make a database of fuzzy rules. Prior to convert the crisp set into fuzzy set, a fuzzy membership function is applied in fuzzy set theory and this function has significant impact on the fuzzy output. Hence in this work, authors are adopted PSO algorithm to tune the value of fuzzy membership function. From results, it is noted that the proposed fuzzy system is more capable to diagnosis of coronary artery disease in comprasion to other methods being compared. Seera and Lim have described a fuzzy min-max neural network for medical data classification task [49]. The proposed system comprises of Fuzzy Min–Max neural network, Regression Tree, and the Random Forest algorithms. The efficacy of the proposed system is tested on Breast Cancer, Diabetes and Liver Disorders diseases and provides significant results. To achieve better accuracy and diagnostic rate of breast cancer affected patients, Ubeyli et al., have proposed an adaptive neuro-fuzzy based inference system (ANFIS) for breast cancer detection [50]. In this work, first order Sugeno model is applied with two fuzzy if-then rules and further, this model is integrated with neuralk network architecture for effective detection of breast cancer diesease. It is found that the ANFIS achieves 99.08 % accuracy rate. A survey paper on application of neuro fuzzy systems is reported in [51]. This

**Table 6**   Comparison of rule based and case based system.

| Criteria | Rule based system | Case based system |
| --- | --- | --- |
| Knowledge representation | Knowledge is represented in the rules | Knowledge is represented in the exemplars |
| Knowledge base set up | Difficult to capture knowledge about a problem with a set of rules | Relatively easy to set up a knowledge base |
| Problem domain | For solving well understood problems | Efficient in problem areas that are not easily interpreted |
| Expansion facility | Hard to expand | Easily be extended by adding new appropriate examples to the set of results |

**Fig. 4** Categories of clinical
support system
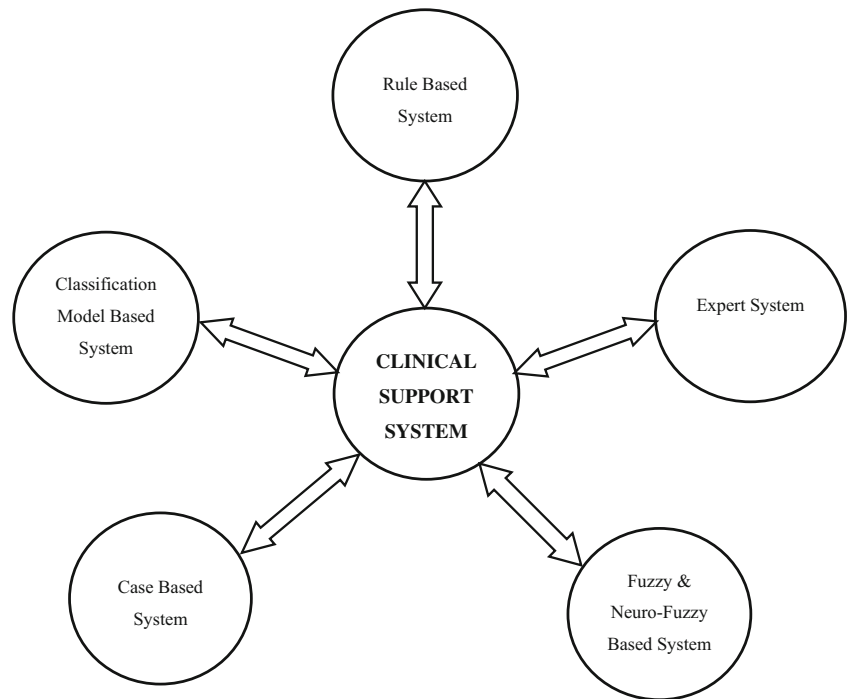


paper also includes the application of neuro fuzzy systems
(NFS) in medical systems. A small portion of dieseases diag-
nosis is also highlighted in above mentioned article.
Papageorgiou et al., have investigated the application of aug-
mented fuzzy cognitive maps to design a novel framework for
dynamic decision support system (DDSS) for medical infor-
matics [52]. In the proposed framework, fuzzy rule extraction
methods are integrated with fuzzy cognitive maps. It is seen

that the proposed DSSS have capability to handle variety of
data and knowledge from various sources and also provides
more accurate results in comprasion to other existing decision
support systems. For pre-determination the risk of cardiovas-
cular disease, Sanza et al., have developed fuzzy rule-based
classification systems using fuzzy rule base and interval-
valued fuzzy sets [53]. In their work, authors initially alter
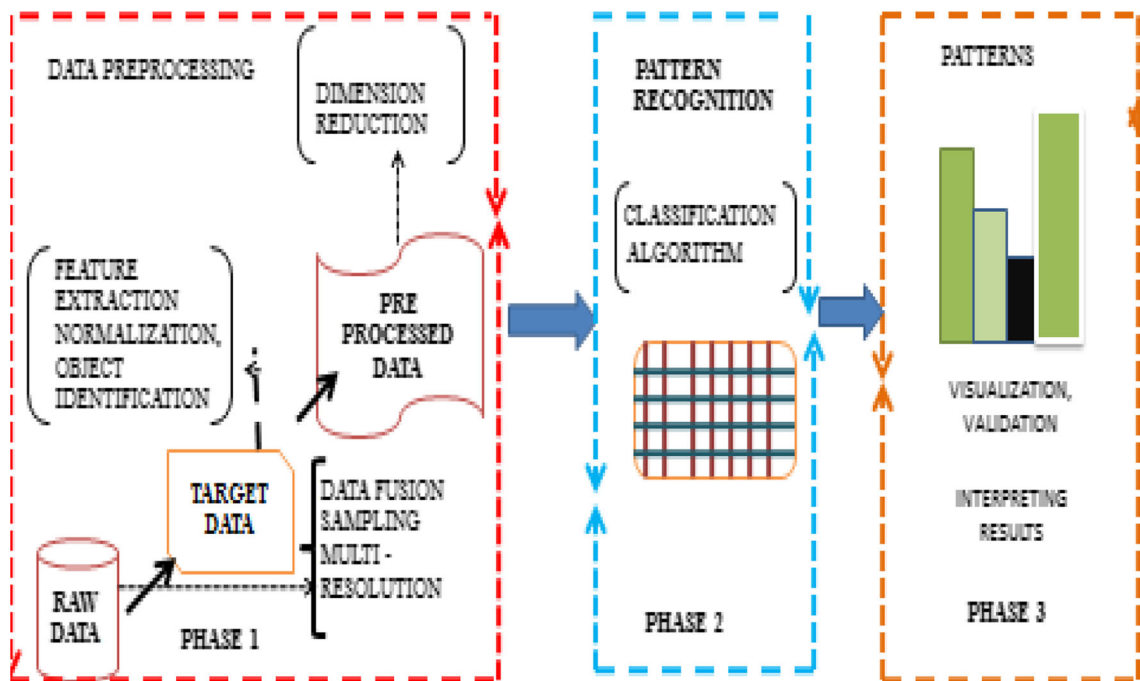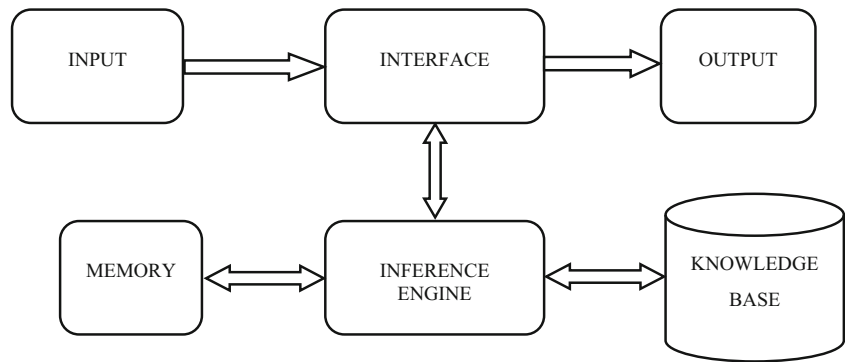the linguistic labels of classifier using interval-valued fuzzy



**Fig. 5** Classification model

Fig. 6 Block diagram of expert system



sets. After alteration process, an operator ($K_a$) is introduced to determine ignorance degree in fuzzy inference process. Further, the genetic algorithm is applied to find the optimized value for parameter of operator ($K_a$) and this operator is applied within each rule of classification systems. It is observed that the proposed fuzzy rule-based classification system is a suitable tool to medical diagnosis of cardiovascular diseases. In continuation of their work, Ubeyli et al., have also proposed an adaptive neuro-fuzzy based inference system for detection of six types of erythemato-squamous diseases [54]. The proposed system obtains higher accuracy rate in comparison to neural network. Kannathal et al., have also shown the applicability of neuro fuzzy network for assessing and examining the different activities regarding autonomic nervous system (ANS) [55]. In their work, authors consider heart signals to find various heart abnormalities. From results, it can be stated that proposed neuro fuzzy tool is an effective and efficient tool for diagnose the different heart abnormalities.

**Rule-based system**

For the knowledge representation, rules are one of the most popular paradigms. A rule based expert system consists of a knowledge base that takes the domain knowledge coded in some patterns. The architecture of rule based system is described in Fig. 8. The utility of rule based systems are mentioned in Table 5. Some of key advantages of rule based systems are given as:

- Modularity: To allow encapsulation of knowledge and expansion of the expert scheme.
- Explanatory: Simply by keeping track of the rules that are fired, an explanation facility can show a range of reasoning resulting to a certain end.
- Human process resemblance: Rules help to explicate the structure of knowledge to the experts.

A considerable amount of research has already been performed in the area of medical sciences using the rule based technique; we have tried reviewing some of the literature work applying this methodology till date. The basic model for rule based systems can be seen in Fig. 7.

Ciabattoniet A., et al. [58] examined three mathematical frameworks introduced by the authors in [59, 60] to decipher the numbers and the inference mechanism based on fuzzy logic probability theory and poroblasitic logic, respectively using Computer-Assisted DIAGnosis (CADIAG-2). The comprehensive precision of proposed diagnosis for CADIAG-2, in comparison to the actual diagnosis is approximately around 80 %. Barakat et al. have presented a SVM based system for diagnosis of diabetes [61]. In the proposed system, rules are extracted using blak box model and the diagnostic rate of proposed system is 94 %. For effective diagnosis of Alzheimer's disease, an association rule based model is reported in [62]. The proposed system evaluates the rules based on the association mining. Further, it is also noticed that in the proposed system, PCA is applied to find the significant features of Alzheimer's disease and these features are further classified by using the SVM.

Fig. 7 Fuzzy based system
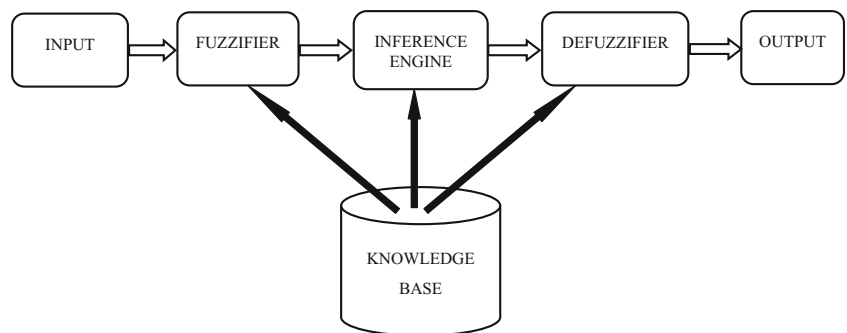
Results indicate that proposed system obtains an accuracy rate up to 91.7 %.

Ell et al., [63] have examined patients with either focal Basal Ganglia (BG) lesions or Parkinson's Disease (PD) on the basis of claims of selective attention and working memory. Focal BG lesions individuals were impaired on tasks with high working memory demand and behave alike to healthy controls on the task with high selective-attention demand. People with PD were also impaired on both projects. The data indicated that the patients with BG lesions exhibited an early-training impairment on a rule-based task in which the demands on working memory were high, but not on a rule-based task that required to selectively attend to one dimension. Instead, it was observed that the PD impairment on the task with high working memory demand was fuelled up by the economic consumption of suboptimal decision strategies.

To reduce the possibility of errors in decision, Astrom and Koker have presented a rule-based parallel neural network for effective diagnosis of Parkinson's Disease (PD) affected patients. The proposed system contains two phases. In first phase, parallel feed-forward neural networks is implemented for classification of PD data. Whereas, the rulebase is designed based on voting decision scheme. Results show that the classification rate of Parkinson's Disease is increased by 8 % in comparison to other method being compared.

Kong et al., have developed a clinical decision support system (CDSS) for observing the stratification of risk in cardic chest pain affected patients [65]. In proposed CDSS, belief rule-based inference concept is adopted for obtaining the better and viable results. The belief rule-based inference is implemented using belief rule base learning method. It is also seen that the proposed CDSS can handle both of clinical domain knowledge and clinical data uncertainties in effective manner. Kumar et al., have described an alternative approach to deal with the problems viz. high complexity, changing medical conditions and low experienced new staff [66]. The proposed approach is the mixture of case-based reasoning and

rule based method and it is totally different from pure rule based reasoning method. Experiment results prove the existence of the proposed method to deal the ICU data.

To investigate the various factors responsible for heart disease in men and women, Nahar et al., have described association rule mining based system to identify the heart disease factors in patients [73]. The proposed system includes the advantage of both the association rule mining and the computational intelligence. Further, it is observed that the rulebase is constructed using apriori, predictive apriori and tertius rule generation algorithms. Results indicate that proposed system is a significant tool for heart disease prediction. Lisboa et al., have examined breast cancer survival data using rule based method [67]. In this work, the analysis of breast cancer data is done in three distinct stages- time-to-event modeling, risk stratification and model interpretation. Event modeling is achieved by selecting the benchmark linear model. Whereas, risk stratification can be measure using cox regression and partially logistic ANN. An automatic rulebase are used to interpret or analysis the breast cancer data with linear model. In their work, authors adopt the Orthogonal Search Rule Extraction (OSRE) method to generate the rule automatically. Further, authors also claimed that the proposed model classifies the patients with low and high risk of breast cancer survivability. For clinical data extraction process, Mykowiecka et al., have presented rule-based information extraction system [68]. The mammography reports and hospital records of diabetic patients are used to achieve the same. The results show that proposed information extraction system reliable and useful for automated clinical data processing. To identify and resolve discrepancies regarding PD affected patients, rule based category learning is reported in [69]. The rule based learning considers various factors like pattern generation, maintenance, shifting and selection for effective treatment of PD patients. Further, neuroimaging and computational modeling are also integrated with rule based learning to get more accurate results. For monitoring heart patients, Seto et al., have developed a rule-based expert system to predict the heart failure symptoms [70]. The proposed rule based expert system is incorporated in mobile phone. In proposed system, rules are designed with the help of heart failure clinician's experts. The performance of the rule base is assessed through a randomized controlled trial. It is found that the proposed rule based expert system improves the clinical management, self-care and quality of life. For effective treatment of knee arthritis, Wei et al., have described a rule based system based on rough set theory [71]. In the proposed system, rules are extracted using rough set theory and specialist opinions. The proposed system consists of three stages. In first stage, data preprocessing is done with the help of specialist opinions. In second stage, two level feature selection methods are adopted to find out relevant features of knee arthritis. In third stage, rules are generated and discretized using rough set theory. It is indicated that proposed system provides valuable
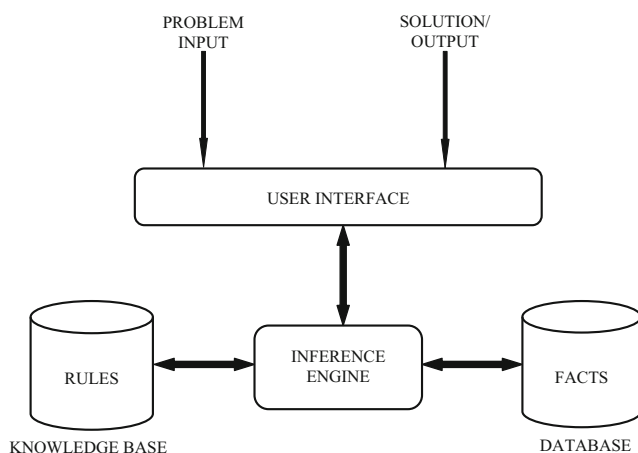


**Fig. 8** Rule based expert system model

and robust results for of knee arthritis. For diagnosing type 2 diabetes, Sevastianov et al., have presented a rule based framework with evidential reasoning for type 2 diabetic patients [72]. The proposed framework is also pinpointing the two collateral problems such as interval-valued belief structures from different sources and comparison of belief intervals. The first problem is handled using interval extended zero method. While, Dumpster–Shafer theory is applied to overcome the belief intervals comparison problem. Results demonstrate that the proposed framework improves the classification rate of type 2 diabetic.

## Case based reasoning (CBR) system

It is a novel AI problem solving paradigm, and has drawn the attention of both soft computing and data mining communities [82]. For a long time, research focused on rule-based expert systems which use production rules to capture domain knowledge and applied an inference algorithm to induce. In recent years, the focal point of research shifted from this principle-based approach to new approaches to knowledge-based reasoning. One of these new approaches is Case-based reasoning. The architecture of case based reasoning system is given in Fig. 9. The advantages of case-based reasoning over reasoning with rules are summarized in Table 6. It has been applied successfully in various medical domains [83–87] and is considered an important instance of decision support systems [88]. It is one of the reasoning methodologies that simulate human reasoning using past experiences to resolve novel problems. Table 7 describes the summary of CBR systems.

Generally, the classical CBR model consists of four step cycle of problem solving [89]:

- RETRIEVE: To call back similar cases to the new case.
- REUSE/ADAPT: To reuse previous results of the most similar example of the new case. Adaptation process can sometimes be included as well to adapt the settlement of the retrieved case is to fit the new case.
- REVISE: To revise the suggested solution for confirmation.
- RETAIN: To retain the learned cases.

For medical data classification, Fan et al., have presented a hybrid model based on fuzzy decision tree and case-based reasoning methods, called it decision making system [90]. In the proposed system, case based reasoning method is employed for data preprocessing phase. Whereas, fuzzy decision tree and genetic algorithm are applied to construct rules for the proposed decision making system. Further, it is noticed that the proposed system cam make more accurate decision in comparison to other models being compared. To diagnosis the chronic obstructive pulmonary disease, Guessoum et al., have developed a case-based decision support system, named it RESPIDIAG [91]. The proposed RESPIDIAG is the combination of case

based reasoning and experts opinions. Five data imputation methods are adopted to handle the missing data. Out of five, three methods compute the similar data for missing information. Results indicate that proposed system provides satisfactory performance. Marling et al., have explained four synergistic systems based on case based reasoning and AI techniques [92]. These are CARE-PARTNER, 4DSS, RHENE and Malardalen Stress System. CARE-PARTNER is a decision support system that is developed for oncology patients, undergone for stem cell transplantation. 4DSS is developed to help the diabetes-1 patients. The proposed system detects problems related to blood glucose control and also gives the valuable suggestions to overcome these problems. RHENE is presented to handle the patients of ERSD disease. Malardalen Stress System (MSS) is developed for the diagnosis and treatment of stress. For effective treatment and also reducing prediction errors, Huang et al., have proposed case based reasoning classifier [93]. The performance of the proposed classifier is compared with PSO optimized ANN and adaptive neuro fuzzy inference system. It can see that the proposed classifier provides better performance in comparison to PSO-ANN and ANFIS. To reduce the threats of chronic disease, Haung et al., have suggested a model for prognosis and diagnosis of chronic diseases [94]. The proposed model incorporates the data mining techniques in case based reasoning. In the proposed model, data mining technique is adopted to devise the rules for examine the health of patients and further, these rules are used to find the presence of chronic diseases. Whereas, CBR is applied for effective diagnosis and treatment of the chronic diseases. Results show that the proposed model is very helpful for doctor as well as patients for the treatment of chronic diseases. Lin and Chuang have presented an intelligent liver diagnosis model for the presence of liver disease in patients [95]. The proposed diagnosis model integrates CBR, ANN and analytic hierarchy process to predict the liver disease affected patients. Further, it is noticed that the proposed is also capable to diagnosis various types of liver diseases. To increases the accuracy rate and also improve the transparency in terms of disease diagnosis, McSherry has developed a conversational case-based reasoning (CCBR) model [96]. In the proposed model, iNN(K)-L method is applied to select the relevant features and further, CCBR algorithm is implemented to diagnosis the disease. Results demonstrate that the existence of the proposed model in the field of medical diagnosis. A case-based reasoning system is also reported for diagnosis of liver disease [97]. The proposed system integrates back-propagation neural network (BPN), regression tree, logistic regression and discriminatory analysis for early diagnosis of liver disease and also enhances classification rate. Another applicability of cased based reasoning is reported in [98]. In the proposed system, GA is applied to obtain appropriate feature of hypertension patients. Further, CBR technique is adopted for the detection of hypertension affected patients. Same with CBR systems, a very few techniques are reported i.e. ANN. An

evidence based decision support system for personal health index is reported in [102].

## Conclusion

The main aim of this study is to identify, categories, and assess the role of different soft computing techniques for diseases diagnosis and prediction. Ninety Two research articles are examined and also arranged these articles into five different classes on the basis of information and data. The main findings of this review are summarized as follows:

- (RQ1) Lot of soft computing techniques have been presented for diseases diagnosis and prediction including SVM, ANN, decision tree, CART, fuzzy, neuro fuzzy and many more. In this work, these techniques are classified into five different classes such as classification model, expert system, fuzzy and neuro-fuzzy based model, rule based model and cased based reasoning approach. It is found that in classification model, SVM and ANN are widely adopted for diseases diagnosis and prediction and both of techniques give viable results. Further, it is also noticed that several studies have also been reported on the parameters optimization of SVM and ANN and observed that PSO and GA algorithm are applied to obtain the optimized value of SVM and ANN parameters. In expert system, rules are obtained with help of fuzzy logic and neuro fuzzy methods. The rules are described in the form of if then structure and it can see that k nearest neighbor and PCA/LDA methods are adopted to find the significant parameters of diseases. The fuzzy and neuro-fuzzy models investigate the applicability of fuzzy logic and neuro fuzzy concept with other soft computing technique except rule

generation. Fuzzy logic and neuro fuzzy concepts have been widely used for parameters optimization of different techniques via AIS, PSO, k-NN etc. In rule based system, rule base have been designed for prediction and diagnosis of diseases. It is noticed that decision tree and some learning methods are adopted for designing the rules. The case base reasoning have also been incorporated with various algorithms like ANN, GA, PSO, AIS for better accuracy and prediction.

- (RQ2) For classification model, PSO-SVM, DT-BN-ANN, GA-ANN and K-Means-SVM have achieved higher accuracy for various types of diseases i.e. 98.91 %, 99.48, 98.24 and 97.38 respectively. While, fuzzy k-nearest neighbor method have obtained an accuracy rate of 96.07 % for expert system class. In fuzzy and neuro-fuzzy based system, k-NN-Fuzzy-AIS have obtained 99.1 % accuracy for diagnosing the disease. In case of CBR system, BPNN-LR-CBR has achieved 95 % accuracy while, DT-GA has obtained 97.44 % accuracy for disease diagnosis and prediction. Overall, it is concluded that SVM and ANN based techniques provide more accurate prediction of diseases rather than other approaches.

- (RQ3) In classification model based system, it is noticed that most of researchers have applied SVM, ANN and tree based techniques for diseases diagnosis. Further, numbers of studies have been presented on optimizing the parameters of SVM and ANN and mostly statistical method like PCA, LDA etc. have been adopted by researchers for optimizing parameter. A very few have been reported on GA and PSO. Hence, it can conclude that optimized values of parameters enhance the performance of techniques. On the other side, it is reported that evolutionary computational methods have always advantage over traditional methods,



Fig. 9 Case based reasoning system
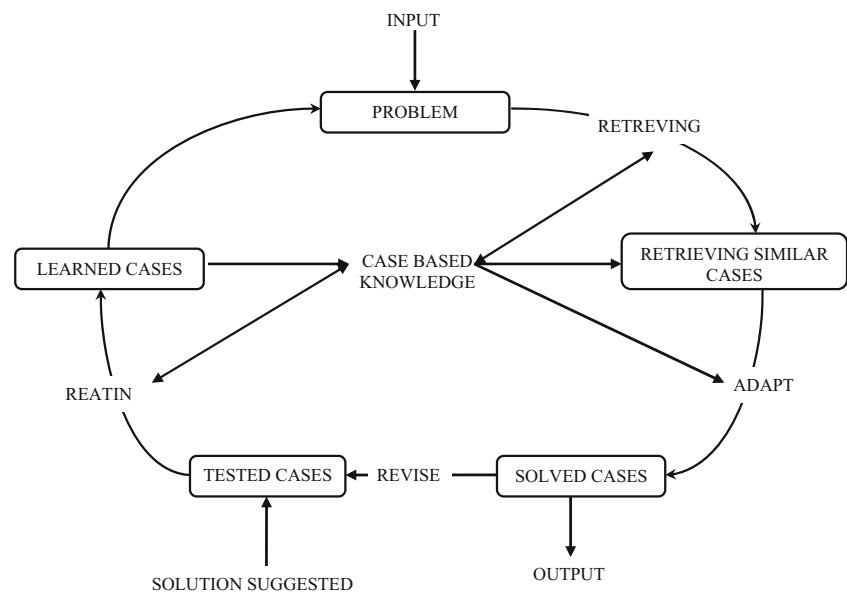
so other evolutionary algorithms can apply for optimizing the parameters of SVM and ANN. It is also observed that DT can work efficiently with linearly separable data and in case of linearly non separable data, its performance very poor. Whereas, SVM works well with linearly non separable data. It is found that for some diseases, SVM works better and for other, DT provides better results. The other issues related to SVM performance is optimal selection kernel, decreased accuracy rate in multi class classification, extensive memory requirement. Hence, to make an effective and better decision support system, it is necessary the proposed technique can work efficiently with both of linearly and non-linearly separable data of diseases. In terms of ANN technique, the numbers of hidden layers and numbers of input neurons affects the accuracy rate. The main disadvantage related to ANN is over fitting and under fitting of data. Another, issues related to ANN is its black box nature and computationally extensive. In expert system approaches, a very few number of soft computing algorithms have been applied for rule generation process like AIS and GA etc. But, large numbers of evolutionary and soft computing methods have been reported in literature and these methods provide better results for solving large number of problems such as ABC, ACO, TLBO, etc. In case of fuzzy and neuro fuzzy based approaches, parameter optimization is one of the major concern area and also affects the performance of the system. The optimized values for parameters also work better than user defines values and provide more accurate results and prediction. Another issue related to applicability of the fuzzy logic either it is used for rule designing or fuzzified the values of parameters. In rule based systems, rules are designed for the diagnosis and prediction of diseases. But, the selection criteria of diseases symptoms are one of main concern for effective rule designing process. So, in rule designing process, feature selection techniques play an important role, if features are not selected correctly then it affects the accuracy rate as well as diagnosis rate. In case of CBR systems, a very few techniques such as ANN, DT and PSO, are integrated to diagnose the diseases.

- (RQ4) Thirty eight research articles have presented the framework and models for effective diagnosis and prediction of various diseases. It is also seen that twenty four models have obtained the accuracy rate higher than 95 %. Hence, these models will help for the doctors and clinicians for effective treatment of patients and also for early detection of diseases. Further, it is also stated that the improvement in clinical practice is laid in the development of computerized systems oriented to diseases symptoms. Rests of articles have described either the new classification algorithm for better diagnose of diseases or presented the new methods for the selection of significant symptoms of diseases.

**Table 7**  Summary of medical data mining techniques using Case Based Reasoning Approach

| Author | Techniques | Disease | Utility/Accuracy |
|---|---|---|---|
| Fan et al. [90] | Combining case based reasoning with fuzzy tree | Breast Cancer | Almost 99.5 % accuracy rate in breast cancer and around 85 % in liver disorder |
| Guessoum et al. [91] | Case based reasoning | Chronic obstructive pulmonary disease | CBR online performs best |
| Marling et al. [92] | Case based reasoning and AI techniques | Stem cell transplantation, diabetes-1, ERSD and Stress | Four synergistic systems are developed which are CARE-PARTNER, 4DSS, RHENE and Malardalen Stress System |
| Huang, et al. [93] | Case based reasoning using PSO based ANN and ANFIS | Breast Cancer | With CBR-based classification 83.60 % accuracy and 91.10 % and 92.80 % with PSO-based ANN classifier and ANFIS |
| Ping et al. [94] | Multiple Measurements Case-Based Reasoning | Liver Cancer | Provides higher accuracy than SingleCBR |
| Lin and Chuang [95] | CBR, ANN and analytic hierarchy process | Liver Disease | Intelligent liver diagnosis model |
| McSherry [96] | iNN(K)-L and CCBR | Lymphography and SPECT heart Diseases | provide accuracy levels of 86.5 % and 84.3 % |
| Chuang [97] | back-propagation neural network (BPN), regression tree, logistic regression and CBR | Liver disease | BPN–CBR method provides significant results in terms of diagnostic accuracy rate, sensitivity and specificity of 95 %, 98 % and 94 %. Respectively |
| Hsu et al. [98] | CBR and GA | Hypertension Disease | Obtained better classification accuracy over other approaches |
| Dina et al. [99] | Case-based reasoning, rule-based reasoning and | Thyroid and Breast Cancer Diseases | Achieved higher accuracy over breast cancer diagnosis and thyroid diseases diagnosis systems |

Apart from these, the several points can be highlighted from this study:

- Twenty six papers are based on unique diseases which show the significant applicability of soft computing techniques for diseases diagnosis and prediction.
- A lot of articles have been reported for effective and better treatment and diagnosis of Cancer and heart diseases.
- Both of the clinical and environmental parameters are considered for diagnosis of various diseases.
- SVM, ANN and FL techniques are widely used in the health care domain for diagnosis as well as earlier identification of diseases symptoms.
- In present scenario, Case based reasoning systems act as an effective tool in medical diagnosis.
- SVM technique obtains better diagnosis accuracy and prediction rate in comparison to other methods being compared.
- The inbuilt parameters of techniques also affect its performance in terms of accuracy, sensitivity and specificity.
- The parameters optimization is also a warm area of research to enhance the performance of techniques and number of algorithm are applied for this.
- PSO algorithm provides better results in terms of parameter optimization among all other methods.
- A very few article reported on the application of evolutionary, swarm and nature inspired algorithms in the field of medical data analysis.

## References

1. Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R., *Advances in knowledge discovery and data mining*. AAAI Press/MIT Press, Boston, 1996.
2. Cios, K.J., Teresinska, A., Konieczna, S., Potocka, J., and Sharma, S., Diagnosing myocardial perfusion SPECT bull's-eye maps-a knowledge discovery approach. *IEEE Eng Med Biol*. 19(4):17–25, 2000.
3. Cios K.J., Moore G.W., Medical data mining and knowledge discovery: an overview. Heidelberg: Springer, pp 1-16, 2000.
4. Cios, K.J., and Kurgan, L.A., *Trends in data mining and knowledge discovery*. Springer, 2002.
5. Cios, K.J., Pedrycz, W., and Swiniarski, R., *Mining methods for knowledge discovery*. Kluwer Academic Publishers, 1998.
6. Fernandez, A., Jesus, M.J., and Herrera, F., On the influence of an adaptive inference system in fuzzy rule based classification systems for imbalanced data-sets. *Expert Syst. Appl*. 36:9805–9812, 2009.
7. Chang, P.C., and Liao, T.W., Combing SOM and fuzzy rule base for flow time prediction in semiconductor manufacturing factory. *Appl. Soft Comput.* 6(2):198–206, 2006.
8. Pendharkar, P.C., Rodger, J.A., Yaverbaum, G.J., Herman, N., and Benner, M., Association, statistical, mathematical and neural approaches for mining breast cancer patterns. *Expert Syst. Appl.* 17: 223–232, 1999.
9. Lai, R.K., Fan, C.Y., Huang, W.H., and Chang, P.C., Evolving and clustering fuzzy decision tree for financial time series data forecasting. *Expert Syst. Appl.* 3:3761–3773, 2009.
10. Giri, D., Acharya, U.R., Martis, R.J., Sree, S.V., Lim, T.C., Ahamed, T., and Suri, J., Automated diagnosis of coronary artery disease affected patients using LDA, PCA, ICA and discrete wavelet transform. *Knowl.-Based Syst.* 37:274–282, 2013.
11. Babaoglu, İ., Findik, O., and Ülker, E., A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine. *Expert Syst. Appl.* 37(4): 3177–3183, 2010.
12. Patil, B.M., Joshi, R.C., and Toshniwal, D., Hybrid prediction model for type-2 diabetic patients. *Expert Syst. Appl.* 37(12): 8102–8108, 2010.
13. Çalişir, D., and Dogantekin, E., A new intelligent hepatitis diagnosis system: PCA–LSSVM. *Expert Syst. Appl.* 38(8):10705–10708, 2011.
14. Sweilam, N.H., Tharwat, A.A., and Moniem, N.A., Support vector machine for diagnosis cancer disease: a comparative study. *Egypt. Inform. J.* 11(2):81–92, 2010.
15. Abdi, J.M., and Giveki, D., Automatic detection of erythematosquamous diseases using PSO-SVM based on association rules. *Eng. Appl. Artif. Intell.* 26(p):603–608, 2013.
16. Bhardwaj, A., and Tiwari, A., Breast cancer diagnosis using genetically optimized neural network model. *Expert Syst. Appl.*:1–15, 2015.
17. Fei, S.W., Diagnostic study on arrhythmia cordis based on particle swarm optimization-based support vector machine. *Expert Syst. Appl.* 37(10):6748–6752, 2010.
18. Chen, H.L., Liu, D.Y., Yang, B., Liu, J., and Wang, G., A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis. *Expert Syst. Appl.* 38(9):11796–11803, 2011.
19. Alkım, E., Gürbüz, E., and Kılıç, E., A fast and adaptive automated disease diagnosis method with an innovative neural network model. *Neural Netw.* 33:88–96, 2012.
20. López, M., Ramírez, J., Górriz, J.M., Álvarez, I., Salas-Gonzalez, D., Segovia, F., Chaves, R., Padilla, P., Gómez-Río, M., and Initiative, A.'s.D.N., Principal component analysis-based techniques and supervised classification schemes for the early detection of Alzheimer's disease. *Neurocomputing.* 74(8): 1260–1271, 2011.
21. Park, Y.J., Chun, S.H., and Kim, B.C., Cost-sensitive case-based reasoning using a genetic algorithm: application to medical diagnosis. *Artif. Intell. Med.* 51(2):133–145, 2011.
22. Latifoğlu, F., Polat, K., Kara, S., and Güneş, S., Medical diagnosis of atherosclerosis from carotid artery Doppler signals using principal component analysis (PCA), k-NN based weighting preprocessing and artificial immune recognition system (AIRS). *J. Biomed. Inform.* 41(1):15–23, 2008.
23. Zheng, B., Yoon, S.W., and Lam, S.S., Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. *Expert Syst. Appl.* 41(4): 1476–1482, 2014.
24. Yeh, D.Y., Cheng, C.H., and Chen, Y.W., A predictive model for cerebrovascular disease using data mining. *Expert Syst. Appl.* 38(7):8970–8977, 2011.
25. Eom, J.H., Kim, S.C., and Zhang, B.T., AptaCDSS-E: a classifier ensemble-based clinical decision support system for cardiovascular disease level prediction. *Expert Syst. Appl.* 34(4): 2465–2479, 2008.
26. Übeyli, E.D., and Doğdu, E., Automatic detection of erythematosquamous diseases using k-means clustering. *J. Med. Syst.* 34(2): 179–184, 2010.

27. Ozcift, A., and Gulten, A., Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms. *Comput. Methods Prog. Biomed.* 104(3):443–451, 2011.

28. Jerez, J.M., Molina, I., García-Laencina, P.J., Alba, E., Ribelles, N., Martín, M., and Franco, L., Missing data imputation using statistical and machine learning methods in a real breast cancer problem. *Artif. Intell. Med.* 50(2):105–115, 2010.

29. Das, R., and Sengur, A., Evaluation of ensemble methods for diagnosing of valvular heart disease. *Expert Syst. Appl.* 37(7): 5110–5115, 2010.

30. Sartakhti, J.S., Zangooei, M.H., and Mozafari, K., Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA). *Comput. Methods Prog. Biomed.* 108(2):570–579, 2012.

31. Nahar, J., Imam, T., Tickle, K.S., and Chen, Y.P.P., Computational intelligence for heart disease diagnosis: a medical knowledge driven approach. *Expert Syst. Appl.* 40(1):96–104, 2013.

32. Morra, J.H., Tu, Z., Apostolova, L.G., Green, A.E., Toga, A.W. and Thompson, P.M., Comparison of AdaBoost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation. *IEEE Trans. Med. Imaging* 29(1), p. 30, 2010

33. Er, O., Yumusak, N., and Temurtas, F., Chest diseases diagnosis using artificial neural networks. *Expert Syst. Appl.* 37(12):7648–7655, 2010.

34. Er, O., Temurtas, F., and Tanrıkulu, A.Ç., Tuberculosis disease diagnosis using artificial neural networks. *J. Med. Syst.* 34(3): 299–302, 2010.

35. Kampouraki, A., Manis, G., and Nikou, C., Heartbeat time series classification with support vector machines. *IEEE Trans. Inf. Technol. Biomed.* 13(4):512–518, 2009.

36. Elveren, E., and Yumuşak, N., Tuberculosis disease diagnosis using artificial neural network trained with genetic algorithm. *J. Med. Syst.* 35(3):329–332, 2011.

37. Cho, B.H., Yu, H., Lee, J., Chee, Y.J., Kim, I.Y., and Kim, S.I., Nonlinear support vector machine visualization for risk factor analysis using nomograms and localized radial basis function kernels. *IEEE Trans. Inf. Technol. Biomed.* 12(2):247–256, 2008.

38. Lee, C.S., and Wang, M.H., A fuzzy expert system for diabetes decision support application. *IEEE Trans. Syst. Man Cybernatics.* 41(1):139–153, 2011.

39. Sengur, A., An expert system based on principal component analysis, artificial immune system and fuzzy k-NN for diagnosis of valvular heart diseases. *Comput. Biol. Med.* 38:329–338, 2008.

40. Sengur, A., An expert system based on linear discriminant analysis and adaptive neuro-fuzzy inference system to diagnosis heart valve diseases. *Expert Syst. Appl.* 35:214–222, 2008.

41. Arsene, O., Dumitrache, I., and Mihu, I., Expert system for medicine diagnosis using software agents. *Expert Syst. Appl.* 42:1825–1834, 2015.

42. Pal, D., Mandana, K.M., Pal, S., Sarkar, D., and Chakraborty, C., Fuzzy expert system approach for coronary artery disease screening using clinical parameters. *Knowl.-Based Syst.* 36:162–174, 2012.

43. Hariharana, M., Polat, K., and Sindhu, R., A new hybrid intelligent system for accuratedetection of Parkinson's disease. *Comput. Methods Med.* 113:904–913, 2014.

44. Keles, A., ESTDD: expert system for thyroid diseases diagnosis. *Expert Syst. Appl.* 34:242–246, 2008.

45. Chen, H.L., Huang, C.C., Yu, X.G., Xu, X., Sun, X., Wang, G., and Wang, S.J., An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach. *Expert Syst. Appl.* 40:263–271, 2013.

46. Ucar, T., Karahoca, A., and Karahoca, D., Tuberculosis disease diagnosis by using adaptive neuro fuzzy inference system and rough sets. *Neural Comput. & Applic.* 23:471–483, 2013.

47. Uguz, H., Adaptive neuro-fuzzy inference system for diagnosis of the heart valve diseases using wavelet transform with entropy. *Neural Comput. & Applic.* 21:1617–1628, 2012.

48. Muthukaruppan, S., and Er, M.J., A hybrid particle swarm optimization based fuzzy expert system for the diagnosis of coronary artery disease. *Expert Syst. Appl.* 39:11657–11665, 2012.

49. Seera, M., and Lim, C.P., Hybrid intelligent system for medical data classification. *Expert Syst. Appl.* 41:2239–2249, 2014.

50. Ubeyli, E.D., Adaptive neuro-fuzzy inference Systems for Automatic Detection of breast cancer. *J. Med. Syst.* 33:353–358, 2009.

51. Kar, S., Das, S., and Ghosh, P.K., Applications of neuro fuzzy systems: a brief review and future outline. *Appl. Soft Comput.* 15:243–259, 2014.

52. Papageorgiou, E.I., A new methodology for decisions in medical informatics using fuzzy cognitive maps based on fuzzy rule-extraction techniques. *Appl. Soft Comput.* 11:500–513, 2011.

53. Sanz, J.A., Galar, M., Jurio, A., Brugos, A., Pagola, M., and Bustince, H., Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system. *Appl. Soft Comput.* 20:103–111, 2014.

54. Pal, D., Mandana, K.M., Pal, S., Sarkar, D., and Chakraborty, C., Fuzzy expert system approach for coronary artery disease screening using clinical parameters. *Knowl.-Based Syst.* 36:162–174, 2012.

55. Kannathal N., Lim, C.M., Acharya, U.R., Sadasivan, P.K., Cardiac state diagnosis using adaptive neuro-fuzzy technique. *Med. Eng. Phys.* 28, pp 809–815.

56. Khatibi, V., and Montazer, G.A., A fuzzy-evidential hybrid inference engine for coronary heart disease risk assessment. *Expert Syst. Appl.* 37(12):8536–8542, 2010.

57. Uzoka, F.M.E., Osuji, J., and Obot, O., Clinical decision support system (DSS) in the diagnosis of malaria: a case comparison of two soft computing methodologies. *Expert Syst. Appl.* 38(3): 1537–1553, 2011.

58. Ciabattoni, A., Picado, M.D., Vetterlein, T., and El-Zekey, M., Formal approaches to rule-based systems in medicine: the case of CADIAG-2. *Int. J. Approx. Reason.* 54:132–148, 2013.

59. Ciabattoni, A., and Vetterlein, T., On the (fuzzy) logical content of CADIAG-2. *Fuzzy Sets Syst.* 161:1941–1958, 2010.

60. Muino, D. P., A probabilistic interpretation of the medical expert system CADIAG-2. *Soft. Comput.* 15:2013–2020, 2011.

61. Pal, D., Mandana, K. M., Pal, S., Sarkar, D., and Chakraborty, C., Fuzzy expert system approach for coronary artery disease screening using clinical parameters. *Knowl.-Based Syst.* 36:162–174, 2012.

62. Chaves, R., Ramírez, J., Górriz, J.M., Puntonet, C.G., and Alzheimer's Disease Neuroimaging Initiative, Association rule-based feature selection method for Alzheimer's disease diagnosis. *Expert Syst. Appl.* 39(14):11766–11774, 2012.

63. Ell, S.W., Weinstein, A., and Ivry, R.B., Rule-based categorization deficits in focal basal ganglia lesion and Parkinson's disease patients. *Neuropsychologia.* 48:2974–2986, 2010.

64. Astrom, F., and Koker, R., A parallel neural network approach to prediction of Parkinson's disease. *Expert Syst. Appl.* 38:12470–12474, 2011.

65. Kong, G., Xu, D.-U., Body, R., Yang, J.-B., Jones, K.M., and Carley, S., A belief rule-based decision support system for clinical risk assessment of cardiac chest pain. *Eur. J. Oper. Res.* 219:564–573, 2012.

66. Kumar, A.K., Singh, Y., and Sanyal, S., Hybrid approach using case-based reasoning and rule-based reasoning for domain independent clinical decision support in ICU. *Expert Syst. Appl.* 36:65–71, 2009.

67. Lisboa, P.J., Etchells, T.A., Jarman, I.H., Hane Aung, M.S., Chabaud, S., Bachelot, T., Perol, D., Gargi, T., Bourdes, V., Bonnevay, S., and Negrier, S., Time-to-event analysis with

artificial neural networks: An integrated analytical and rule-based study for breast cancer. *Neural Netw.* 21:414–426, 2008.

68. Mykowiecka, A., Marciniak, M., and Kupsc, A., Rule-based information extraction from patients' clinical data. *J. Biomed. Inform.* 42:923–936, 2009.

69. Price, A., Filoteo, J.V., and Todd, M.W., Rule-based category learning in patients with Parkinson's disease. *Neuropsychologia.* 47:1213–1226, 2009.

70. Seto, E., Leonard, K.J., Cafazzo, J.A., Barnsley, J., Masino, C., and Ross, H.J., Developing healthcare rule-based expert systems: case study of a heart failure telemonitoring system. *Int. J. Med. Inform.* 81:556–565, 2012.

71. Wei, M.H., Cheng, C.H., and Li, J.Y., Discovering medical resource utilization in total knee arthroplasty (TKA) using rule-based method. *Arch. Gerontol. Geriatr.* 55:157–164, 2012.

72. Sevastianov, P., Dymova, L., and Bartosiewicz, P., A framework for rule-base evidential reasoning in the interval setting applied to diagnosing type 2 diabetes. *Expert Syst. Appl.* 39(4):4190–4200, 2012.

73. Nahar, J., Imam, T., Tickle, K.S., and Chen, Y.P.P., Association rule mining to detect factors which contribute to heart disease in males and females. *Expert Syst. Appl.* 40(4):1086–1093, 2013.

74. Jung, H., Yang, J., Woo, J.I., Lee, B.M., Ouyang, J., Chung, K., and Lee, Y., Evolutionary rule decision using similarity based associative chronic disease patients. *Clust. Comput.* 18(1):279–291, 2015.

75. Toro, C., Sanchez, E., Carrasco, E., Mancilla-Amaya, L., Sanín, C., Szczerbicki, E., Graña, M., Bonachela, P., Parra, C., Bueno, G., and Guijarro, F., Using set of experience knowledge structure to extend a rule set of clinical decision support system for alzheimer's disease diagnosis. *Cybern. Syst.* 43(2):81–95, 2012.

76. Chen, R.C., Huang, Y.H., Bau, C.T., and Chen, S.M., A recommendation system based on domain ontology and SWRL for anti-diabetic drugs selection. *Expert Syst. Appl.* 39(4):3995–4006, 2012.

77. Stoean, R., and Stoean, C., Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection. *Expert Syst. Appl.* 40(7):2677–2686, 2013.

78. Sarkar, B.K., Sana, S.S., and Chaudhuri, K., A genetic algorithm-based rule extraction system. *Appl. Soft Comput.* 12(1):238–254, 2012.

79. Ang, J.H., Tan, K.C., and Mamun, A.A., An evolutionary memetic algorithm for rule extraction. *Expert Syst. Appl.* 37(2):1302–1315, 2010.

80. Mohamed, M.H., Rules extraction from constructively trained neural networks based on genetic algorithms. *Neurocomputing.* 74(17):3180–3192, 2011.

81. Kumar, Y., and Sahoo, G., Prediction of different types of liver diseases using rule based classification model. *Technol. Health Care.* 21(5):417–432, 2013.

82. Sharaf-El-Deen D.A., Moawad I.F., Khalifa M.E., A new hybrid case-based reasoning approach for medical diagnosis systems. J. Med. Syst. 38:9, Springer, 2014.

83. Ocampo, E., Maceiras, M., Herrera, S., Maurente, C., Rodríguez, D., and Sicilia, M.A., Comparing Bayesian inference and case-based reasoning as support techniques in the diagnosis of acute bacterial meningitis. *Expert Syst. Appl.* 38(8):10343–10354, 2011.

84. Begum, S., Ahmed, M.U., Funk, P., Xiong, N., and Folke, M., Case-based reasoning systems in the health sciences: a survey of recent trends and developments. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 41(4):421–434, 2011.

85. Huang, M.-L., Hung, Y.-H., Lee, W.-M., Li, R.K., and Wang, T.-H., Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *J. Med. Syst.* 36(2):407–414, 2012.

86. Teodorović, D., Šelmić, M., and Mijatović-Teodorović, L., Combining case-based reasoning with bee Colony ptimization for dose planning in well differentiated thyroid cancer treatment. *Expert Syst. Appl.* 40(6):2147–2155, 2013.

87. Depeursinge, A., Vargas, A., Gaillard, F., Platon, A., Geissbuhler, A., Poletti, P.-A., and Müller, H., Case-based lung image categorization and retrieval for interstitial lung diseases: clinical workflows. *Int. J. Comput. Assist. Radiol. Surg.* 7(1):97–110, 2012.

88. Bichindaritz, I., and Montani, S., Advances in case-based reasoning in the health sciences. *Artif. Intell. Med.* 51:75–79, 2011.

89. Chattopadhyay, S., Banerjee, S., Rabhi, F.A., and Acharya, U.R., A Case-Based Reasoning system for complex medical diagnosis. *Expert. Syst.* 30(1):12–20, 2013.

90. Fana, C.Y., Chang, P.C., Lin, J.J., and Hsieh, J.C., A hybrid model combining case-based reasoning and fuzzy decision tree for medical data classification. *Appl. Soft Comput.* 11:632–644, 2011.

91. Guessoum, S., Tayeb, L.M., and Lieber, J., RESPIDIAG: a case-based reasoning system for the diagnosis of chronic obstructive pulmonary disease. *Expert Syst. Appl.* 41:267–273, 2014.

92. Marling, C., Montani, S., Bichindaritz, I., and Funk, P., Synergistic case-based reasoning in medical domains. *Expert Syst. Appl.* 41:249–259, 2014.

93. Huang, M.L., Hung, Y.H., Lee, W.M., Li, R.K., and Wang, T.H., Usage of case-based reasoning, neural network and adaptive neuro-fuzzy inference system classification techniques in breast cancer dataset classification diagnosis. *J. Med. Syst.* 36(2):407–414, 2012.

94. Ping, X.-O., Tseng, Y.-J., Lin, Y.-P., Chiu, H.-J., Lai, F., Liang, J.-D., Huang, G.-T., and Yang, P.-M., A multiple measurements case-based reasoning method for predicting recurrent status of liver cancer patients. *Comput. Ind.* 69:12–21, 2015.

95. Lin, R.H., and Chuang, C.L., A hybrid diagnosis model for determining the types of the liver disease. *Comput. Biol. Med.* 40(7):665–670, 2010.

96. McSherry, D., Conversational case-based reasoning in medical decision making. *Artif. Intell. Med.* 52(2):59–66, 2011.

97. Chuang, C.L., Case-based reasoning support for liver disease diagnosis. *Artif. Intell. Med.* 53(1):15–23, 2011.

98. Hsu, K.H., Chiu, C., Chiu, N.H., Lee, P.C., Chiu, W.K., Liu, T.H., and Hwang, C.J., A case-based classifier for hypertension detection. *Knowl.-Based Syst.* 24(1):33–39, 2011.

99. Sharaf-El-Deen, D.A., Moawad, I.F., and Khalifa, M.E., A new hybrid case-based reasoning approach for medical diagnosis systems. *J. Med. Syst.* 38(2):1–11, 2014.

100. Long, N.C., Meesad, P., and Unger, H., A highly accurate firefly based algorithm for heart disease prediction. *Expert Syst. Appl.* 42(21):8221–8231, 2015.

101. Kumar, S.U., and Inbarani, H.H., Neighborhood rough set based ECG signal classification for diagnosis of cardiac diseases. *Soft. Comput.*:1–13, 2016.

102. Chen, L., Li, X., Yang, Y., Kurniawati, H., Sheng, Q.Z., Hu, H.Y., and Huang, N., Personal health indexing based on medical examinations: a data mining approach. *Decis. Support. Syst.* 81:54–65, 2016.

103. Prasad, V., Rao, T.S., and Babu, M.S.P., Thyroid disease diagnosis via hybrid architecture composing rough data sets theory and machine learning algorithms. *Soft. Comput.* 20(3):1179–1189, 2016.