

Computational decision support system in healthcare: a review and analysis

Ravita Chahar*

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

Received: 09-December-2020; Revised: 20-February-2021; Accepted: 24-February-2021

©2021 Ravita Chahar. This is an open access article distributed under the Creative Commons Attribution (CC BY) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

A decision support system (DSS) may help to synchronize comprehensive computational aspects of computational problems like knowledge discovery, processing, and pattern visualization. It provides an effective channel for decision making by churning huge datasets. Primarily, DSS has been deployed in domains like business process management, health informatics, and even for managing smart devices. Numerous algorithms have been proposed to augment the efficacy of DSS and are undergoing refinement. Most recently, DSS has been trialed in the healthcare industry because there is a need for an intelligent decision support system. It may be helpful in the process of automation along with the solution generation and evaluation. In this paper, broad analysis and discussion on the applicability and suitability of the methods related to DSS and algorithms have been discussed along with the role of information and communications technology (ICT). It also includes problem-based discussion and suggested solutions along with different cases in the healthcare system. It covers the data analysis with the domain intelligence impact along with the information processing for knowledge extraction and discovery. It also covers some of the clinical decision aspects for understanding the impact of other correlated medical resource systems. The methodological and computational analysis includes data preprocessing, knowledge extraction, interpretation, decision-making model, and the influencing factors in the performance analysis. The main data mining methods considered here for the DSS system in case of healthcare informatics discussion were association rule mining, clustering, classification and optimization algorithms. The machine learning aspects covered here in three ways supervised, semi-supervised, and unsupervised for the decision analysis based on the healthcare system. Finally, based on the methodological and computational applicability different decision-making scenarios have been discussed and analyzed for the analysis of the combination and nature of applicability. Our study and analysis provide an analytical and computational perspective in terms of the health care system, influencing parameters, their applicability, a methodological perspective, decision-making process, traditional methods, and the challenges along with the suggested measures for the future. It's also helpful in the process of maintaining the internal and external aspects which is more reliable in performance aspects of DSS.

Keywords

DSS, ICT, Data mining, Machine learning, Soft computing, Evolutionary algorithms.

1.Introduction

In the current scenario, technology is playing a vital role in every field. It makes everyone's life easier, better, and smoother [1–3]. Imagine any stuff in today's era is tough without technology. It helps in the development of communication between anything and everything. It may be adapted to communication, data processing, predication, or healthcare system. With the help of technology, problem detection and decision processing can become faster and easier [2, 3]. Technology helps us to change our life towards a technical world. The data is growing with the same proportion as the population grows.

On this earth, every person is a user who wants to use data either in any form like whether he wants to see it as a problem or he wants to find a solution related to data collected.

The prominent role of information communication technology has an immense impression of the health care system. This system needs development in different areas with varying specifications. This information technology is acceptable and adaptable to each and every one. Although there are numerous challenges which are faced by many health care sectors like keeping the records of patients, maintaining the information of hospitals and patients, information regarding medical equipment, medication records, and many more [4]. All these types of task may be handled easily through a

*Author for correspondence

decision support system (DSS) and information and communications technology (ICT) system [5].

Now a day's health care is the most sensitive topic amongst all. The use of information communication technology in terms of health care can be categorized as follows:

1. Health and Education
2. Hospital data management
3. Health care resources and availability

This division is because so that people can access and communicate with each other in a faster and effective way. It is very useful in knowledge representation and data availability. The analysis is shown below:

- It is very important to have knowledge about the issues related to your health. All these types of data are available on the internet where different users have shared their experiences. Sometimes it may be relevant and sometimes it may not be. But it is concluded that with the help of others, we can easily identify the symptoms and can gain knowledge about precaution and cure. One kind of awareness can be generated from this education. This platform gives us also the opportunity to teach someone from our good and bad experiences.
- It helps in successfully leading any organization. It helps in improving the safety and security of patients by learning the latest technologies. Knowledge of doctors can be enhanced by learning new skills.
- Research is an ongoing process. Now when it is concerned with health it is a never-ending process. Every time someone is doing something different so that he can find all possible measures to reduce the impact of the disease on health. It always inspires every person to explore and give a beneficial conclusion which can help with several people in their different domains.
- A lot of data is increasing day by day now it becomes important to keep that data because of that data only we can create information which will be helpful for several years. It behaves like a repository where all the necessary things are being stored and used when required.

Like DSS and ICT different data mining and machine learning techniques may be helpful in the decision-making system [1]. But it depends upon the knowledge and applicability.

The major challenges in the advanced computing system for health informatics are as follows:

1. There is a need for parametric evaluation and impact determination for efficient computation handling.
2. If the parametric variations are not scalable then how it will be synchronized and combined with another level of methodological evaluation?
3. The association between methods and their threshold ranges should be acquired and enhanced automatically.

The above challenges show the need for the automation strategy in terms of efficient framework development for the process of advance computing applicability with the inclusion of DSS and ICT based systems in the health care industry.

The decision-making process in the DSS system depends on different factors, including internal and external aspects. The internal aspects include the data analysis, patterns impact, associative impacts of attributes, and covering the complete reports on these factors. The external effects may vary depending on the problem statements. In general, it can be specific to the tuning parameters, algorithm constants, and the influencing parameters. The tuning parameters can be used in terms of problem optimization and the influencing parameter can be used in the overall decision making. So, the complete decision is the hybridization of different aspects of DSS. This study has mainly four objectives. Firstly, to explore the DSS system in the healthcare domain with computational variability along with the analysis of the influencing parameters and the role of ICT. The second objective is to survey and analyze the methodological aspects in terms of the applicability of different methods in the DSS system for the healthcare system. The third objective is in the direction of results, discussion based on clustering, classification, and optimization algorithms for the performance comparison in the same domain.

This paper objectives are as follows:

1. To explore the health care industry in terms of data availability, management, and storage.
2. To discuss and analyze different health-related issues with the methodological impact with the outcomes in terms of detection and prediction.
3. To review and analyses the methodological aspect for the prospective of the discussion of the current trends.

4. To identify the advantages along with the gaps for the suggestive explanation based on different measures.
5. To elaborate on the role of DSS and ICT with data mining and machine learning techniques.

This paper covers the following:

1. It covers the motivation, objectives, background of DSS along with the need in the healthcare system. It also covers the complete rationale behind the fulfillment of objectives. Then the exploration regarding the emergence of decision support systems in healthcare.
2. Next, the decision support system has been included in the case of the healthcare system. It covers different methods along with the study and analysis of the impact of individual and combined methods.
3. The suitability of each method with different problem scenarios has been discussed in detail. It also covers the empirical result evaluation based on the current and past trends of the related methods used.
4. Then computing parameters impact in terms of internal and external influencing factors have been discussed. It has been discussed in terms of selection criteria, preprocessing, knowledge discovery, automation, and extraction.
5. Then the combined effects of methods and the parameters used in the healthcare system which can affect the decision processing have been discussed and analyzed with different cases and performance analysis.
6. Finally, based on the results, discussion, and analysis of the future framework have been suggested for the healthcare system. It also includes future suggestions and current limitations.

2.Related work

Ayanso and Lertwachara (2015) [6], discussed ICT development in terms of affordability and adaptability. They have considered supervised and unsupervised techniques. They have explored the development and affordability profiles. Ganser et al. (2015) [7], discussed the telehealth. It has been discussed in terms of remote patient monitoring. It has the possibility of healthcare cost reduction. It has been used for the cost-saving prediction, medical cost variability and in case of intervention patients and controls. Fuller et al. (2016) [8], discussed the anxiety affects in terms of virtual team interactions. They have conducted an elaborated review based on computer-mediated communication (CMC). Data has been gathered based on survey questions, messages,

and scores of team projects. Cai et al. (2016) [9], discussed the beck depression inventory (BDI). They have investigated BDI for the depression prediction. They have considered gender differences as the parameter for the prediction of depression profiles. For the subset selection, they have considered the feature selection method. They have suggested that self-accusation, mood, work inhibition, loss of libido and social withdrawal are the major factors for the depression in males. In the case of females, they have suggested work inhibition, indecisiveness, distortion of body image, crying, mood, and work inhibition. Alabdani et al. (2016) [10], discussed depression and suggested a framework. Their suggested framework is helpful in the detection of depression in the case of Saudi patients. They have used pattern recognition for depression detection. The parameters considered here are eye movements, facial expressions, body movements, head movements, heart rate, and body. Their combinations have been considered also. Bochicchio et al. (2016) [11], studied data analytics in complex healthcare environments in case of big data. They have suggested the applicability of clustering and classification algorithms in the big data framework through the massive amount of data. Dhaka and Johari (2016) [12], discussed the exponential rise in data storage in terms of the health care industry. They have discussed the use of clustering and other data mining techniques. They have also used and considered the deployment of big data in terms of MongoDB in healthcare data. Li et al. (2016) [13], discussed domain-specific problems in case of big databases. They have proposed a hierarchical integration approach. Their approach is highly scalable and appropriate in big data analytics platforms such as Hadoop. Koppad and Kumar (2016) [14], studied decision support system in terms of big data for the customer experience and transformation. Their main goal is to analyze and discuss healthcare analysts. It has been discussed in terms of chronic obstructive pulmonary disease (COPD). They have applied a decision tree algorithm for better performance. Guo et al. (2017) [15], presented linear discriminant analysis (LDA) for the volunteer's categorization. They have proposed a multi-objective particle swarm optimization (MOPSO). It has been used for the detection of depression. They have achieved 100% accuracy. They have considered the minimization and maximization factors along with the distance factors. Chaitanya et al. (2017) [16], discussed anxiety detection and identification of data retrieved based on neurological consideration. They have considered RGB light-emitting diode (LED) and

electroencephalography (EEG)-sensor. It has been used as the Arduino driven platform. The results indicate the strength in terms of performance. Costarides et al. (2017) [17], discussed the anxiety disorder. They have proposed a generalized anxiety disorder intervention method for adolescents. It is helpful in the emotional resilience strength; confidence builds up and skills development through different online games. Hajian and Moradi (2017) [18], discussed the depression. They have suggested that it is a major concern nowadays. They have suggested that the depression level is directly associated with the structural changes in the brain. So, authors have suggested that based on the brain signal, depression level can be detected. They have used a fuzzy classifier and support vector machine (SVM) for the depression quantification. Their results outperform in terms of percent relative difference (PRD). Haritha et al. (2017) [19], discussed the anxiety detection system development. They have considered SVM for the backend classifier. For the validation, they have considered anxiety disorders individuals. Their working process is based on the kernel matching. For the higher dimension linear mapping, they have used fisher vector encoding (FVE). They have also considered principal component analysis (PCA) for the feature vector transformation and the input features mapping. Li et al. (2017) [20], discussed and explored the augmented reality (AR) and virtual reality (VR). They have measured the heart rates and skin conductance. Their results show the preliminary experiments in terms of subjective scales. Mazlan et al. (2017) [21], discussed the learning management system. It shows information access in terms of ICT in the teaching and learning environment. It also provides the impact of social impact and social anxiety information access to find the list in terms of different social and influential factors. Alharbi (2017) [22], discussed the health informatics e-Learning object repository (HiLOR) project. They have presented regarding the design, implementation and usability evaluation. It has been done in terms of the proposed system. Their results show the performance in terms of support personalization. Faruk et al. (2017) [23], discussed the quality of health care access. They have discussed the role of ICT in the case of the health care delivery system. They have discussed the virtual clinics and the integration of health information systems. Chatterjee et al. (2017) [24], discussed the Internet of Things (IoT) in the case of traditional healthcare methods. Their paper main implication is to develop the decision support system based on IoT. It has been useful for

determining cardiovascular disease. They have developed a decision support system for this in terms of efficiently predicting diseases. Ghoshachandra et al. (2017) [25], discussed the increasing rate of the number of patients in case of hypertension. They have discussed a suitable time to measure blood pressure. For this purpose, they have proposed eHealth. It is as a self-care android application. For the data storage, they have used azure SQL database. Their application has six main components. These are managed profile, measure blood pressure, provide knowledge, record treatment, provide reminders, and perform data analysis. The main aim of their proposed app to assist senior citizens. Islam et al. (2017) [26] discussed the heart diseases condition. So, they have recommended that the early detection of such type of disease can help to resolve this effectively. The photoplethysmography (PPG) signal was suggested by the researchers for detecting the heart disease. As compared to traditional study of diagnosis there were the possibility of tissue damage in the treatment of patients. They extracted the features from the facial video of human. They also proposed it as an efficient and cost-effective methods. Pouriye et al. (2017) [27], investigated and analyzed regarding the heart disease prediction. They have considered the Cleveland data set. They have used decision trees, Naïve Bayes, multilayer perceptron, k-nearest neighbor, single conjunctive rule learner, radial basis function and support vector machine. Better results have been obtained in the combination of boosting SVM. Priyanga and Naveen (2017) [28], discussed the significance of data mining techniques in the reference of pattern finding. They have used naïve Bayes. They have considered the parametric indicators as no, low, average, high and very high. They have performed classification and prediction. An accuracy of 86% has been achieved. Yeh et al. (2018) [29] discuss the impact of heart disease. According to the authors, it is a common cause of death worldwide. They have suggested regarding the availability of the data also. They have suggested that the combination of data mining techniques can be helpful and can achieve better prediction results. Lee et al. (2018) [30] discussed the brain activity investigation. It is based on electroencephalography (EEG) videography. Their EEG features may be helpful in driver monitoring and smart cars emotional states detection. Aceto et al. (2018) [31] discussed the progressive view of ICT in the healthcare area. They have correlated the conjunction of healthcare and ICT. For study and analysis, they have considered more than 300 papers. They have surveyed and analyzed the complete

correlation based on the industry paradigm. Liu et al. (2018) [32] discussed the negative emotion or anxiety. They have discussed the various ways of analysis of the motions. They have proposed a multi-modal system for this. Their results show the effectiveness in terms of real-life scenarios. Fritchman et al. (2018) [33], discussed machine learning techniques. They have discussed secure multiparty computation (SMC). It has been discussed in terms of privacy-preserving machine learning (PPML). They have discussed the cryptographic protocols. It has been used for sending the data securely in the service-oriented framework. They have also discussed the deployment protocol along with the cloud-based protocols. Breil et al. (2018) [34], discussed the use of Electronic medical information and health systems. Their analysis is helpful in the specification and evaluation in the case of health information systems along with the management. Khader et al. (2018) [35] discussed the technology acceptance model (TAM). It has been discussed in terms of Jordanian and Palestinian students' attitudes assessments. It has been discussed in terms of health informatics. Kaur and Bawa (2018) [36], discussed data mining for the voluminous data extraction. They have proposed the popularity of data mining techniques in today's era. Data mining shows the improvement in clinical decisions and patients' survival time. Their main objective is to explore and analyses data mining techniques. Lee et al. (2019) [37] discussed public health sustainability and practices in terms of ICT. They have investigated health risk communication and governance regarding public health. They have evaluated the e-government adoption model (GAM). Their results show that social media, citizens perceive and security are a major concern that will influence health sustainability. Kimani (2019) [38] discussed anxiety reduction based on sensor-driven technologies. The

author evaluated the real-time system and explores the behavior also. Heir automated framework is capable of physiological detection. Li (2019) [39] discussed digital simulation technology. Their investigation suggests that multi-touch gestures may improve science education in showing the learning activities and different levels of a digital solution. Ma and Wang (2019) [40] discussed the diagnosis of depression. They have suggested that it solely depends on the behavioral symptoms of the person. They have considered Twitter data. They have extracted it from the data of the huge symptoms from Twitter. They have proposed a hybrid method for this. It is the combination of natural language processing and semantic graph. Amith et al. (2019) [41], discussed about the consumer health vocabulary. They have developed SKOSbased based on consumer health vocabulary dataset. They have suggested providing consumer-level meaning for the semantically linked dataset. They have suggested further expansion in the terminology. Balatsoukas et al. (2019) [42], discussed the process of involving users. It has been discussed in terms of the mobile decision support tool. They have discussed the implications in terms of the design process and their recommendations. Yahyaoui et al. (2019) [43], discussed the medical decision support systems (DSS). They have discussed this in terms of clinical decision making. They have considered a support vector machine (SVM) and random forest (RF) for classification. They have also used a convolutional neural network (CNN) as the deep learning (DL) case. It has been used for the prediction and detection of diabetes patients. Their results show 76.81%, 65.38%, and 83.67% respectively in the case of DL, SVM, and RF. *Table 1* shows the computational result-based analysis.

Table 1 Computational result-based analysis

S.No	Reference	Method	Results	Limitations
1	Vogel, (2016)[44]	ICT in healthcare	The results support the help in education through the use of ICT.	It is specific to children's parameters.
2	Stoicu-Tivadar (2010) [45]	ICT continuity framework for healthcare	A technical model and framework have been presented.	Model should be validated with different parameters.
3	Melzer et al. (2015) [46]	Leakage cancellation	Gain of 6DB has been achieved in terms of sensitivity.	It can be validated with different hardware and other specific requirements.
4	Zeadally and Bello (2019) [47]	IoT in healthcare	Global networks can be provided through the IoT. It can be combined with a different medical system.	They have suggested the need for collaboration with AI techniques.

S.No	Reference	Method	Results	Limitations
5	Yadav et al. (2019) [48]	Virtual reality interfaces	The results support the behavioral interventions.	They have suggested the inclusion of group-specific parameters.
6	Puli and Kushki (2019) [49]	Anxiety Detection in the presence of motion	They have achieved overall arousal detection accuracy of 93%.	They have suggested for the naturalistic setting based validation also.
7	Cacheda et al. (2019) [50]	Random forest algorithm	The detection model shows 10% improvement.	They have suggested validating the different environment also to evaluate the state of the art results.
8	Wang et al. (2018) [51]	ACP approach used in Health care systems.	They have proposed a framework based on the parallel healthcare systems (PHSs) along with the artificial systems with the computation mechanism and parallel execution capability (ACP). It can analyze and evaluate different therapeutic regimens that can be capable of decision-making support and real-time optimization.	They have suggested the extension in terms of different other diseases.
9	Meephak et al. (2018) [52]	Elder eat	The main motive of their project is to prosper and implement an Android platform, namely for the daily food consumption recording. The platform name is elder eat. The main benefit of this platform is its consumption behavior.	It can be extended for different regions and density for the global impact.
10	Navin et al. (2017) [53]	Support epidemic control	Their proposed methodology discussed the health surveillance program. They have applied the pilot testing of the proposed prototype. They have analyzed substantiating cause with some messages for the prevention. It has been sent immediately through the same app.	There is a need for the inclusion of the computational methodology.
11	Movahedi et al. (2018) [54]	Sequential healthcare data	Time-stamped events sequence representation can be presented by electronic health records (EHRs). They have focused on the development of the methodological approach. It has been extracted based on different patterns. The pattern extraction based on the left ventricular assist device (LVAD). It has been discussed based on the advanced heart failure.	There is a need for exploration of experimental results.
12	Malavasi et al. (2019) [55]	IoT based home system for integrated care services	They have discussed the design, development, and implementation plan for the solutions based on IoT for the integrated care services. They have discussed the major target participants that may be disabled people, unpaid carers, paid carers, people who are professionals in the health system and people who are socially active.	They have suggested to include the gamification elements.
13	Xia et al. (2019) [56]	DSS for acute coronary syndrome patient similarity	They have discussed the data-driven clinical decision support system. They have also proposed a framework. The main focus of this project is data-driven CDSS. It has been discussed in terms of acute coronary syndrome patients.	They have suggested the need for privacy-preserving learning and secure aggregation inclusion.
14	Chisengantambu-Winters et al. (2020) [57]	Decision-making dependency model	Their main aim is to explore the decision-making processes for the nurses. They have discussed the small group of nurse managers study. It has been discussed in terms of the rural area. This case considered in the case of South Australia.	Different other cases may be considered and validated.

3. Methods and computational decision making

A DSS with ICT may help to synchronize comprehensive computational aspects of computational problems like knowledge discovery, processing and pattern visualization [58–61]. It provides an effective channel for decision making by churning huge datasets. Primarily, DSS has been deployed in domains like business process management, health informatics and even for managing smart devices. Numerous algorithms have been proposed to augment the efficacy of DSS and are undergoing refinement. Most recently, DSS has been trailed in the healthcare industry because there is a need for an intelligent decision support system. It may be helpful in the process of automation along with the solution generation and evaluation [4]. The decision-making process in the DSS system depends on different factors, including internal and external aspects. The internal aspects include the data analysis, patterns impact, associative impacts of attributes and covering the complete reports on these factors. The external effects may vary depending on the problem statements. In general, it can be specific to the tuning parameters, algorithm constants, and the influencing parameters. The tuning parameters can be used in terms of problem optimization and the influencing parameter can be used in the overall decision making. So, the complete decision is the hybridization of different aspects of DSS [5].

This section mainly contains the four objectives in which the first is to explore the DSS system in the healthcare domain with computational variability with different domains. The second objective is to analyze the algorithm aspects in terms of the applicability of different methods in the DSS system for the healthcare system. The third objective is in the direction of results, discussion based on clustering, classification and optimization algorithms for the performance comparison in the same domain with the deployment. Finally, based on the analysis and discussion some insights on the applicability and suitability of the above-said methods and algorithms have been discussed with the limitations. It also includes problem-based discussion and suggested solutions along with different cases in the healthcare system. It covers the data analysis with the domain intelligence impact along with the information processing for knowledge extraction and discovery. It also covers some of the clinical decision aspects for understanding the impact of other correlated medical resource systems. The methodological and computational analysis includes data pre-processing,

knowledge extraction, interpretation, decision-making model and the influencing factors in the performance analysis.

The suitability of each method with different problem scenarios has been discussed in detail. It also covers the empirical result evaluation based on the current and past trends of the related methods used. Then computing parameters impact ICT in terms of internal and external influencing factors that have been discussed. It has been discussed in terms of selection criteria, pre-processing, knowledge discovery, automation, and extraction. Then the combined effects of methods and the parameters used in the healthcare system which can affect the decision processing have been discussed and analyzed with different cases and performance analysis.

Data mining is a process of knowledge discovery through relevant extracted patterns. In general, the data mining methods can be categorized as classification, clustering, regression, prediction and association rules [62].

In data mining, association rule mining (ARM) plays a critical role [63]. With the help of this rule we can identify the correlations amongst different items or things for various application areas [63]. For finding the set of frequent pattern they have provided the threshold value. It means one the with the help of this value of minimum support we can determine the complete set of patterns. Similarly for the business processes, few support values they have provided with the use of minimum and maximum approaches. Like this make the tedious task to filter it out in several aspects. *Figure 1* shows the key association rules for a specific set. T1, T2, and T3 show the transactions. *Figure 2* shows the complete data analysis process. It shows the pre-processing mechanism, rule discovery procedure, and pattern extraction. *Figure 3* shows the data processing in the computational system (Phase-I). It includes data attribution, data cleaning, data transformation, and scaling. Data attribution deals with attribute selection and synchronization. Missing and unwanted removal has been done in the cleaning phase. Mapping and transformation have been done in the transformation and scaling phase. *Figure 4* shows the data classification or grouping in the computational system (Phase-II). It includes classification and clustering techniques. *Figure 5* shows data classification or grouping performance (Phase-III). It shows the final data transformation, extraction, evaluation, performance measures, and error

evaluation. *Figure 6* shows the complete correlation based on association. *Figure 7* shows the final association based on machine learning. It shows the complete mechanism to learn and predict. This

mechanism shows automate improvements. It shows that accuracy and performance may increase due to the accessibility of mechanism through learning and predictions in case of machine learning.

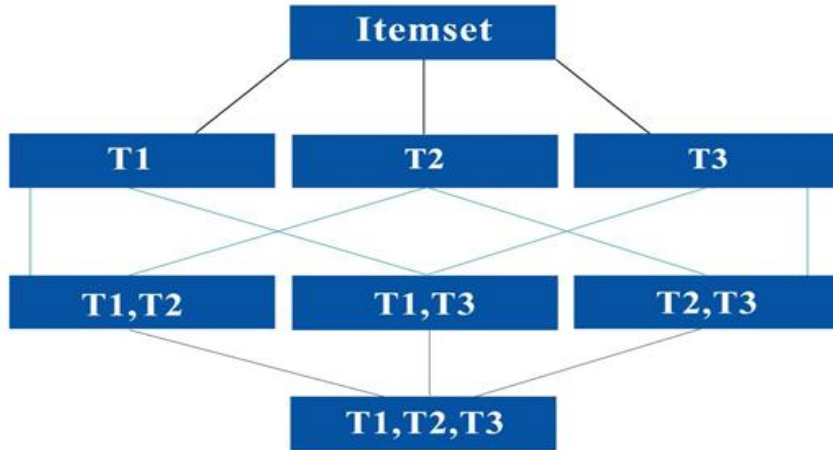


Figure 1 Itemset association

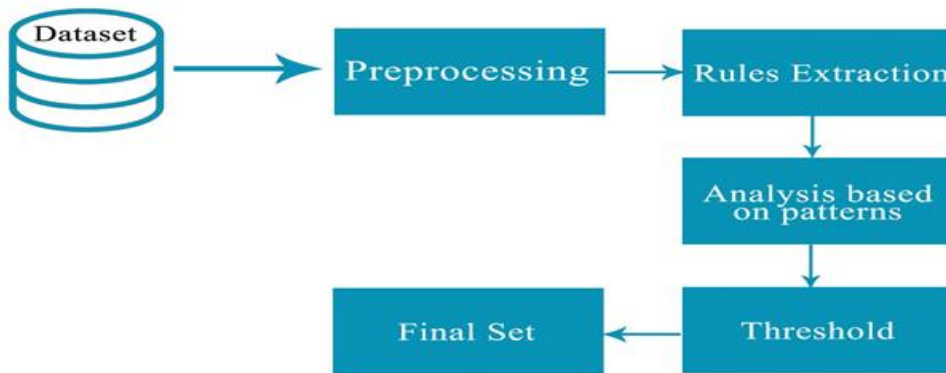


Figure 2 Data analysis procedure

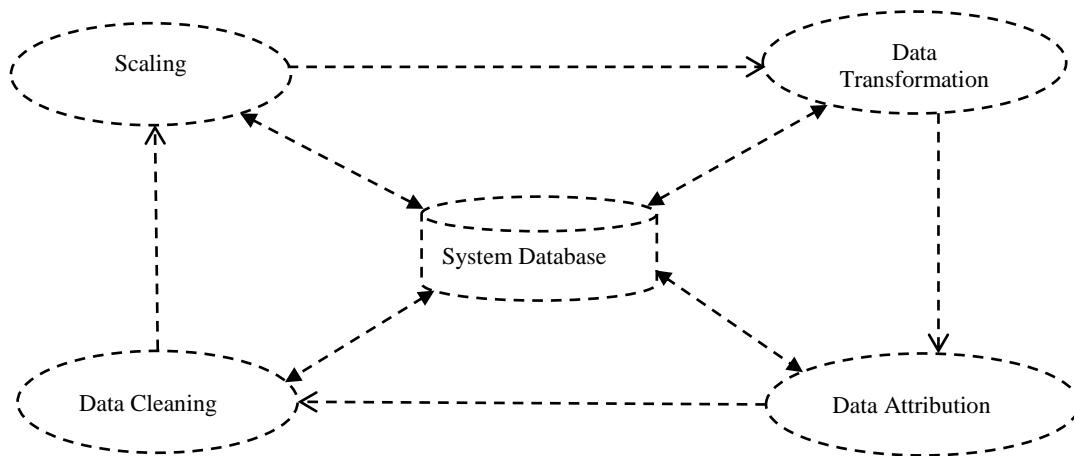


Figure 3 Data processing in the computational system (Phase-I)

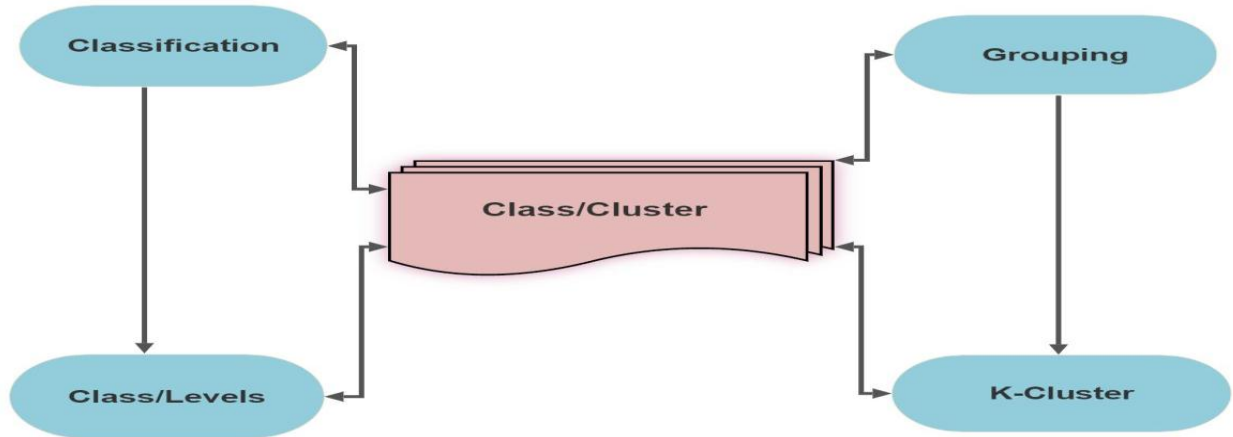


Figure 4 Data classification or grouping in the computational system (Phase-II)

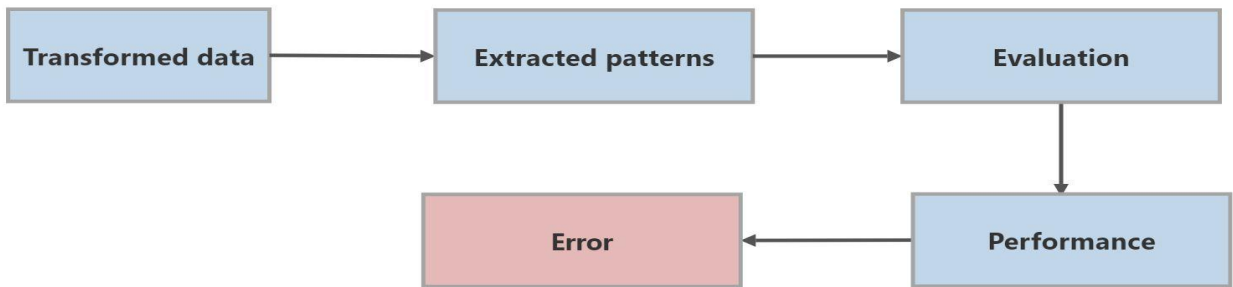


Figure 5 Data classification or grouping performance (Phase-III)

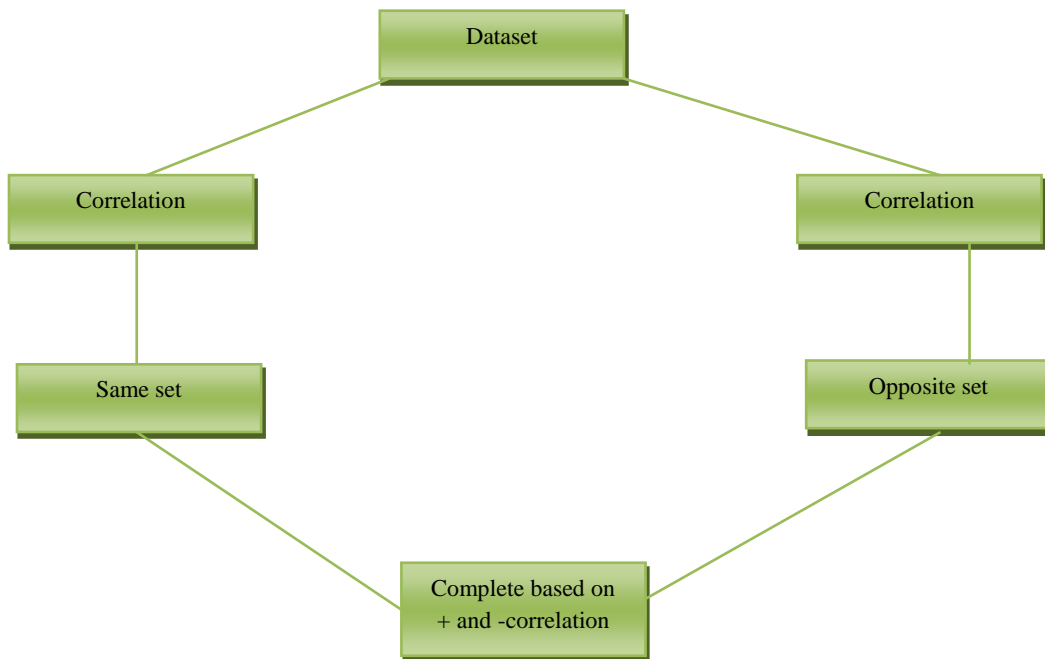


Figure 6 Complete correlation based on association

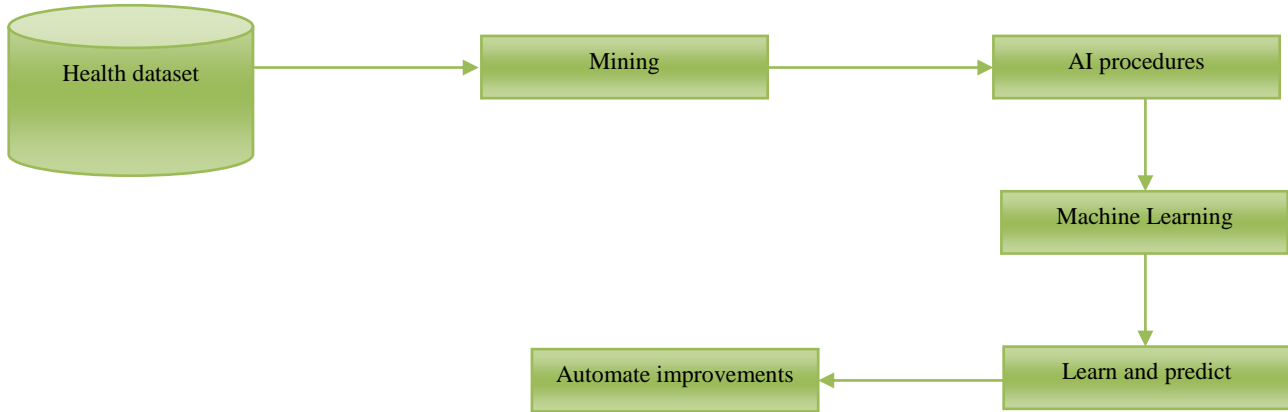


Figure 7 Final association based on machine learning

Clustering algorithms have been used for the grouping of data based on similar properties or the attributes likeness [64]. The major features of the clustering algorithm are feature selection, algorithm selection, cluster validation and finally the result interpretation. Some of the important clustering algorithms are as follows:

3.1 Hierarchical clustering

It is of two types. One is agglomerative and the second is divisive.

3.1.1 Agglomerative hierarchical clustering technique

In this type of clustering technique individual cluster has been considered as the data point. This process is iteratively merged for the formation of the k-cluster. It has been used for the proximity matrix calculation and

the updation of the same. It supports bottom-up design [65].

3.1.2 Divisive hierarchical clustering technique

It follows top to down approach. Here all data points collectively have been used for the formation of a single cluster. In each iteration single cluster has been created [65].

3.2 K-means clustering

This is an unsupervised algorithm of learning and very general in use with wide applicability. In this procedure number of clusters has already been defined [66,67]. In which k-centroids have been defined for each cluster formation in each iteration. It has been processed until the same centroid will not be repeated. The process is shown in *Figure 8*.

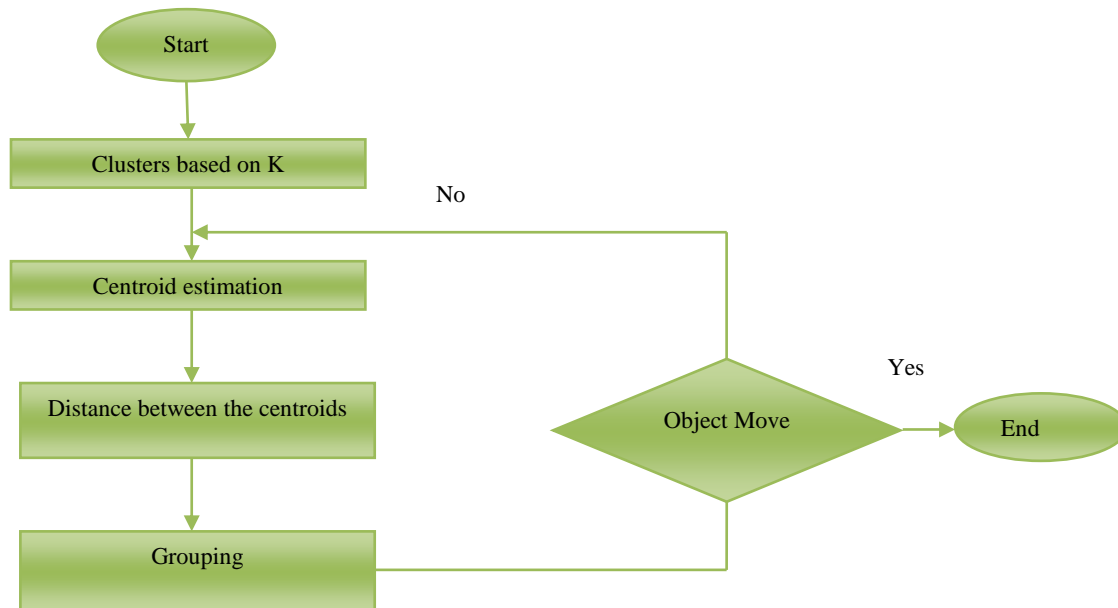


Figure 8 K-means process

Algorithm: K-means algorithm

Step 1: Selection of the dataset based on the weight or the pre-processed data.

Step 2: The iteration has been started with the initial centroid.

Step 3: For each iteration, it has been calculated randomly and the process will be executed based on the unbiased call for every single iteration.

Step 4: Then Euclidean distance has been applied for the similarity measurement

$$X(c) = \sum_{j=1}^k \sum_{i=1}^n ||d_i^{(j)} - c_j||^2$$

$d_i - c_i$ is the Euclidean distance.
 k is the number of clusters.
 n shows the case numbers.

Step 3: The cluster center formation has been created based on the similarity score to the repeated match

$$c_i = \left(\frac{1}{n_i}\right) \sum_{j=1}^{n_i} d_i$$

Step 4: Partitioning has been performed.

Step 6: Final clusters have been performed.

3.3 Fuzzy C-means clustering

It is also used in the extensive domains for the clustering. It has been used in such a manner that the data points can be fit in more than one cluster [68, 69]. The process is shown in *Figure 9*.

Algorithm: Fuzzy c-means algorithm

$d_1, d_2, d_3 \dots, d_n$ forms a D segment for the data point values from a set of $C = \{c_1, c_2, c_3 \dots, c_n\}$.

Step 1: Cluster center selection and assignment.

Step 2: Computation of membership values:

$$U = \sum_{i=1}^c u_{ij} = 1, \forall j = 1, \dots, n$$

Step 3: Distance has been calculated in the following manner with the degree of membership assignment:

$$J(U, c_1, c_2, \dots, c_c) = \sum_{i=1}^c J_i = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2$$

Centroid indicated by c .

d_{ij} denotes the distance.

Step 4: Final data point assignments have been performed.

Step 5: Final clusters have been formed.

3.4 Mountain clustering

In the mountain clustering algorithm, the computations have been calculated based on the exponential growth with the dimensional increase of the data [70].

3.5 Subtractive clustering

In the case of subtractive clustering, each data point has the possibility of the formation of the data center

as the individual in each iteration [71]. It has been done through the surroundings data point density.

3.6 Self-organizing maps

One of the most popular mapping strategies is self-Organizing Maps which are based on the neural network model. This supports the competitive learning network for the assessment without human intervention [72, 73]. It is unsupervised learning.

Now we discussed some of the important technique which has been used in machine learning and may be useful in health informatics domain. Classification algorithms have been considered now in terms of machine learning algorithms for further discussion. These algorithms are capable of learning based on history and the processed data.

Some of the important algorithms are discussed below [74–79].

3.7 Linear regression

Linear regression is the machine learning algorithm for the prediction mechanism. It predicts the values based on the dependent and independent variable. There are two types of variables in linear regression one is criterion variable and another is predictor variable. The variable which we are predict is called criterion variable and other is the variable on the basis for the predication has been performed is called predictor variable. Here y and x are considered as the criterion variable and predictor variable respectively.

Regression can be performed as follows:

$$y = a + bx \tag{1}$$

b is the intercept value.

Consider error e with the intercept value. It can be shown as follows:

$$y = a + bx + e \tag{2}$$

It can be expanded with n number of independent variables.

$$y = a + b1x1 + b2x2 + \dots + bnxk + e \tag{3}$$

The prediction can be shown as follows:

$$y = a + b1x1 + b2x2 + \dots + bnxk \tag{4}$$

It can be express a one variable like the following equation:

$$b = \frac{\sum xy}{\sum x^2} \tag{5}$$

$$a = y - bx \tag{6}$$

It can be express a two variable like the following equation:

$$\text{bone} = \frac{(\sum x_2^2) \sum(x_1 y) - \sum(x_1 x_2) \sum(x_2 y)}{(\sum x_1^2) \sum(x_2^2) - \sum(x_1 x_2)^2} \quad (7)$$

$$\text{btwo} = \frac{(\sum x_2^2) \sum(x_2 y) - \sum(x_1 x_2) \sum(x_1 y)}{(\sum x_1^2) \sum(x_2^2) - \sum(x_1 x_2)^2} \quad (8)$$

3.8 Naive Bayes classifier

This algorithm is based on Bayes’ Theorem. The main assumption in this classifier is the main predictors are completely independent [80]. So, it can be said that the individual presence of any feature

does not rely or depend on any other feature. It has been useful for large datasets.

It is helpful in the calculation of posterior probability. It is denoted by P(c|x). It can be deduced from the prior probability of the class (P(c)), prior probability of the predictor (P(x)) and posterior probability of the class (P(x|c)). It can be represented by the below equation.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)} \quad (9)$$

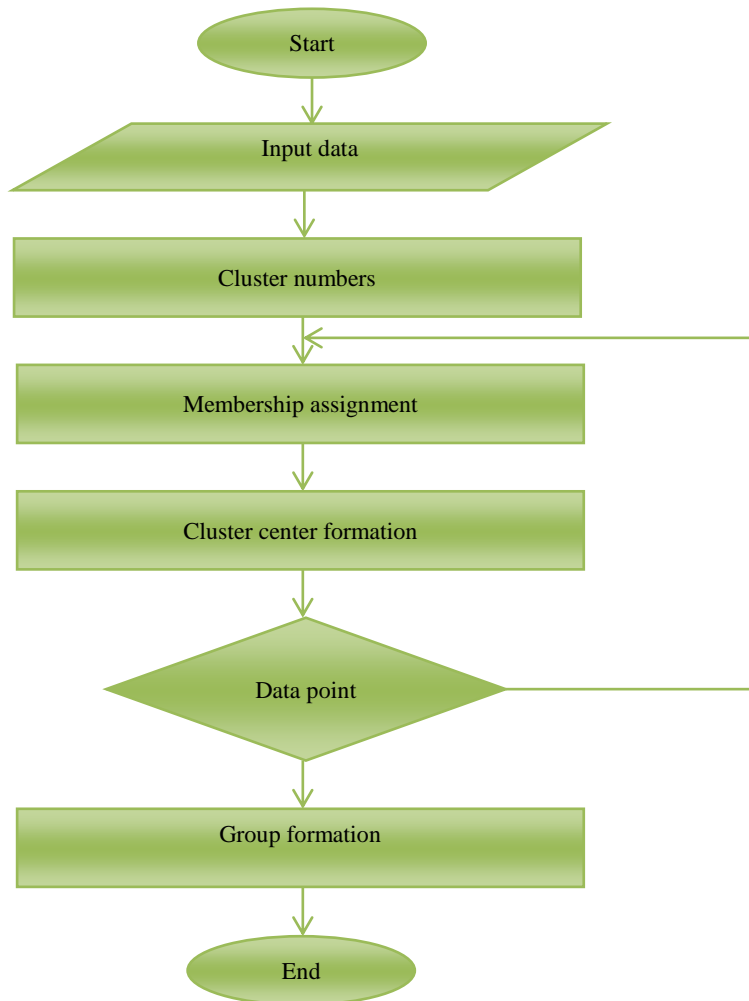


Figure 9 Fuzzy c-means process

3.9 Nearest neighbor algorithm

In this technique, a complete class has been considered for learning and it should be connected with other data points for the same [39]. For further labelling new data points have been considered.

For the K nearest neighbor measurement different distance measures have been used. For K = 1, nearest neighbor has been assigned automatically.

Euclidean distance

$$\sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (10)$$

Manhattan distance

$$\sum_{i=1}^k |x_i - y_i| \quad (11)$$

Minkowski distance

$$[\sum_{i=1}^k (|x_i - y_i|^q)]^{\frac{1}{q}} \quad (12)$$

Hamming distance has been used for the categorical variables. The following equation has been used for the following.

$$DH = \sum_{i=1}^k |x_i - y_i| \quad (13)$$

$x = y \rightarrow D = 0$
 $x \neq y \rightarrow D = 1$

3.10 Support vector machine

It is categorized as the discriminative classifier. It works on the labelled training data. It has been used for the analysis and regression [37].

The input variable space splitting can be presented by the following equation.

$$B0 + (B1 \times X1) + (B2 \times X2) = 0 \quad (14)$$

B1 and B2 shows the coefficient.

X1 and X2 are the input variables.

B0 is the intercept.

The prediction algorithm can be shown as follows:

$$f(x) = B0 + \text{sum}(a_i \times (x, x_i)) \quad (15)$$

There is the need of different kernel for the implementation of the Support vector machine. Kernel is basically a dot product.

$$\text{Kernel}(x, x_i) = \text{sum}(x \times x_i) \quad (16)$$

In case of polynomial kernel, the above equation can be shown as follows:

$$\text{Kernel}(x, x_i) = 1 + \text{sum}(x \times x_i) \quad (17)$$

When Gamma function is used it is called radial kernel. It can be shown as follows:

$$\text{Kernel}(x, x_i) = e^{(-\text{gamma} \times \text{sum}((x - x_i)^2))} \quad (18)$$

3.11 Logistic regression

This algorithm provides a statistical analysis of the data [81]. The outcomes have been calculated based on the independent variables. It is capable of quantitatively factors calculation and classification. It can be represented as sigmoid function:

$$\frac{1}{1 + e^{-x}} \quad (19)$$

It can be represented as the regression equation as follows:

$$p(X) = \frac{e^{(b_0 + b_1 \times X)}}{1 + e^{(b_0 + b_1 \times X)}} \quad (20)$$

$$n \frac{p(X)}{(1-p(X))} = b_0 + b_1 \times X \quad (21)$$

The cost function of logistic regression can be explained as follows:

$$\text{Cost}(h\theta(x), y) = \begin{cases} -\log(h\theta(x)) & \text{if } y = 1 \\ -\log(1 - h\theta(x)) & \text{if } y = 0 \end{cases} \quad (22)$$

It can be explained in terms of probabilistic function. If the probability is 1 then it can be represented as follows:

$$Y = p(Y=1|X) \quad (23)$$

If the probability is 0 then it can be represented as follows:

$$1 - Y = p(Y=0|X) \quad (24)$$

$$P(Y|X) = Y \cdot (1 - Y) \quad (25)$$

$$\text{If } Y=1 \text{ then } P(Y|X) = Y \quad (26)$$

3.12 Decision trees

This classification has been performed based on the tree structure. It has been formed based on the smaller subsets for the association of the decision tree in terms of incremented fashion. So, the results have been calculated based on the tree node and leaf node [53].

The following steps has been performed in case of decision tree:

Step 1: Determine those classes which are mainly important in the decision of the system.

Step 2: The entropy has been calculated for each domain and also for the specific constraints also.

Step 3: Other attributes have been considered also.

$$\text{Entropy}(S, T) = \sum P(C)E(C) \quad (27)$$

Where P shows the probability and E shows the entropy calculation in terms of entropy. S shows the divided part and T is target variable.

So, information gain can be represented as follows:

$$\text{Gain}(T, X) = \text{Entropy}(T) - \text{Entropy}(T, X) \quad (28)$$

Step 4: Then information gain has been calculated for each division.

Step 5: Perform the division for each iteration based on tree-based structure.

3.13 Random forest

It is a group based learning method used for the classification. The working processing relies on a multitude of decision trees. The construction is completely based on a multitude of decision trees

[82]. The results have been classified based on the mode of the classes or based on the regression.

In case of random forest scikit-learn mechanism, has been adopted as shown below:

$$N_{ij} = w_j C_j - w_{left} C_{left(j)} - w_{right} C_{right(j)}$$

Where

N_{ij} : The implication of node j

$w_j C_j$: The importance of weighted number.

$left(j)$: Left side node

$right(j)$ = Right side node

3.14 Neural network

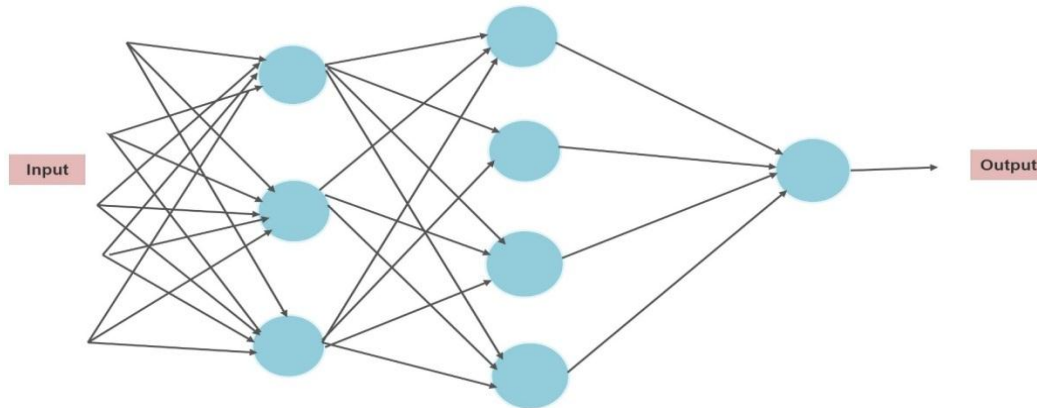


Figure 10 Neural network architecture

There are mainly three phases for the methodology deployment. These are preparing, refine and promote and final deployment. In the preparation phase, the core team has been created. The documents and the algorithms have been reviewed. Then high-level deployment plan has been executed. Finally, the deployment establishment has been performed.

In the next phase refine and promote procedure has been performed. In this process, the deployment plan has been refined. The procedure is then calculated based on the number of approaches for the combination and the same set is used for the hybridization and further stuff for the next cycle and pre-processing has been performed.

In the last phase final deployment has been performed based on different cyclic approaches. It includes initialization, execution, identification, and updation. This process also includes the following:

It consists of neurons. It has been arranged in different layers [83]. The input and outputs are associated with different layers. It has been associated with the unit feed mechanism for the layer-wise traversing and accepting it in several layers for the final processing and computation.

It can be constructed in three types of layers (*Figure 10*).

Input layer: It represents the initial data.

Hidden layers: It shows the intermediate layer between input and output.

Output layer: It shows the results.

- Information Cleaning: Data is removed through different procedures, for example, filling out in missing esteems, smoothing the boisterous information, or fixing the necessary information.
- Information Integration: It assembled the data with various portrayals and check out the clashes staying inside the information. All such types of things are settled here.
- Information Transformation: It makes the data standardization, and then summed it up.
- Information Reduction: It further intends to exhibit a decreased render of the information in an information stockroom.
- Information Discretization: It involves the diminishment of various assessment of a constant property by partitioning it into the scope of characteristic interims.

3.15 Soft computing and optimization techniques in healthcare

Soft computing may provide a better solution with differential variability for the real-life complex problems and health data [84]. It is based on fuzzy logic, genetic algorithms, artificial neural networks

(ANN), machine learning, and expert systems [85–90].

Different classification techniques have been used in a variety of ways with different training data for the targeted output. Soft computing along with the optimization techniques can be useful in a variety of way including classification and clustering [86]. There exist some different optimization techniques including ant colony optimization (ACO), particle swarm optimization (PSO), genetic algorithm (GA), artificial bee colony (ABC) algorithm, and teaching-learning-based optimization (TLBO) algorithm. These algorithms, in general, require three things population size, number of generations and algorithm-specific parameters [87].

PSO has been prospered by Kennedy, Eberhart, and Shi [88]. It is used for the social behavior simulation based on metaheuristic approach [89].

ACO algorithm has been inspired by real ant behavior. It provides the pheromone-based

communication for the searching mechanism for the food source. It has been developed by Marco Dorigo in 1992 [91]. It is efficient in model-based search [92, 76].

TLBO algorithms deal with the teaching-learning aspects [93]. Rao and Patel [93] suggested that all the nature-inspired algorithms require controlling parameters. These are population size and iterations in terms of generators They have also suggested that the other algorithms also require some other tuning parameters specific to the algorithms [93]. They have suggested that these affect the performance aspects also negatively sometimes. They have suggested that TLBO may perform better as it does not require these types of algorithm-specific parameters.

Figure 11 shows the hybrid components of soft computing and optimization techniques with computational aspects [94, 95]. Figure 12 shows the soft computing techniques classification [94, 95].

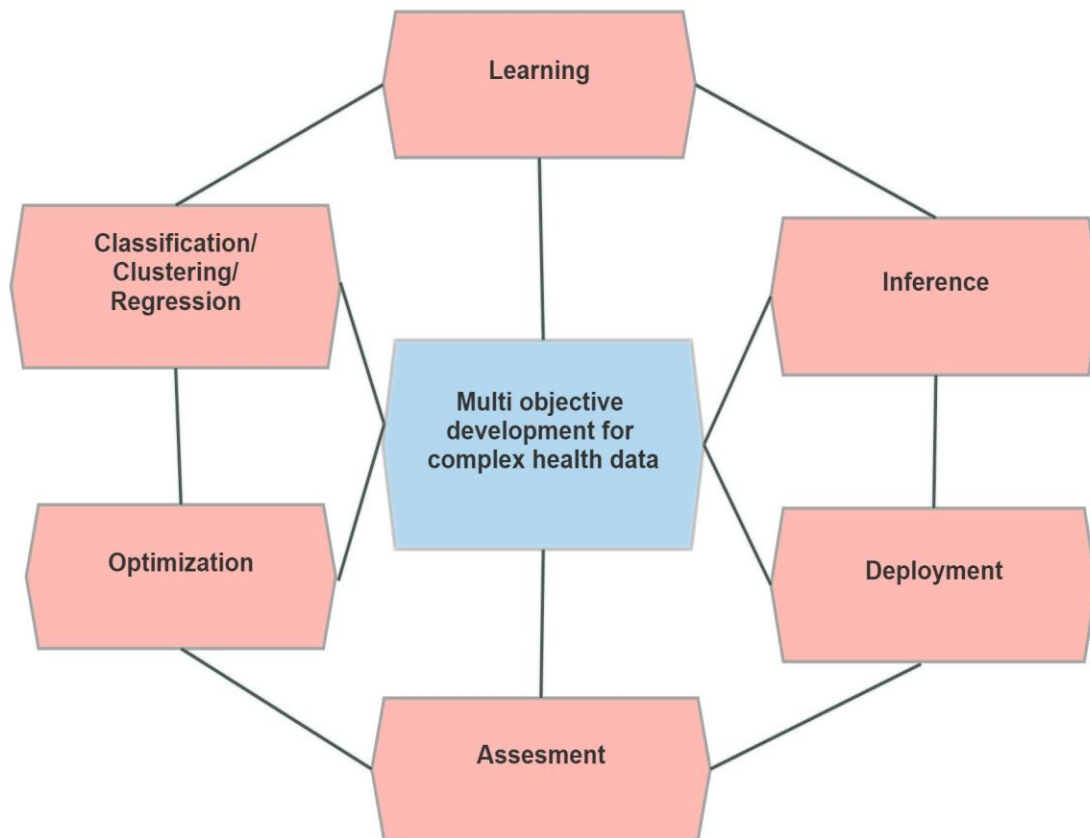


Figure 11 Hybrid components of soft computing and optimization techniques with computational aspects

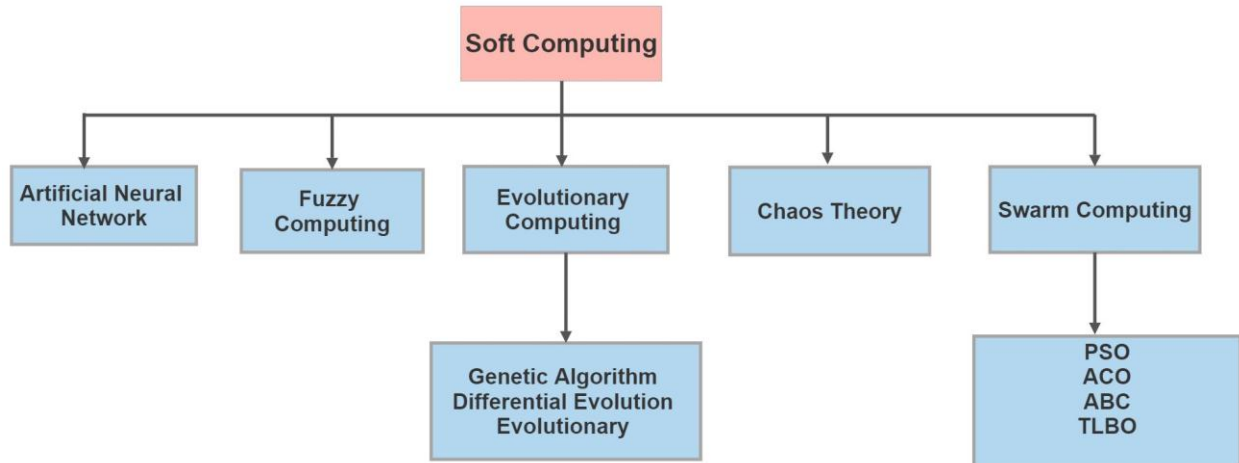


Figure 12 Soft computing techniques classification

4. Results and discussion

The results are discussed in terms of their applicability for different methods and algorithms. These results are based on the study presented and the literature discussed. Some of the literature is considered for the cases related to health-data which was used for the detection and prediction process.

It has been found that data mining algorithms are helpful in the case of initial pre-processing, data extraction and rule finding (*Figure 13*). It also shows the relationship between different methods of use and adaptability. It is based on the review of various papers for different parameters.

We found that there are some clinical applications also exists which can help in the treatment of patients. With the help of these applications doctors can analyze the data (history) and that can help in taking the decision. In conjunction with the related clinical research, data mining is used for conducting the examination for treatment. Further, it helped in formulating the process of diagnosis and improving the quality of treatment. The line of treatment on the basis of decision support system can be based on knowledge or non-knowledge based system. In the knowledge-based system they use some inference mechanism on the data repository system whereas in non-knowledge-based system they applied artificial intelligence network. In non-knowledge based system they analyzed the patterns and then determine the relationship among them. The symptoms they found can be the reason for further treatment. When it is compared to traditional decision support system it was difficult to make suggestions or we can say that this decision support of clinical basis needs the

software designed for providing direct relation to their problems.

The attribute values may be the determining factor in the case of health-related data. So, the determination of each attribute individually along with the methodological view should be considered differently. As we know the structure of biological data is different. The methodological analysis includes data pre-processing, knowledge extraction, interpretation, decision-making model, and the influencing factors in the performance analysis. For this purpose, association rule mining, clustering, classification, and nature-inspired algorithms can be used. The machine learning aspects may be helpful in automated learning and for accuracy improvement. This computational analysis provides an analytical and computational perspective in terms of the mental health care system, influencing parameters, their applicability, a methodological perspective, decision-making process, traditional methods, and the challenges along with the suggested measures for the future framework. The major algorithms used in the initial stage are data mining, evolutionary algorithms, and machine learning for automation and learning process.

In the case of final accuracy calculation, data mining and evolutionary algorithms are found to be prominent. But the major significance in these approaches is the combinatorial use related to every method (*Figure 14*). It is very important to have knowledge about the issues related to health informatics. It also determines types of data available on the internet where different users have shared their experiences. The effect is completely based upon the idea of utilization and also on the tuning of

algorithm-specific parameters. So, it is hard to design a common framework for all diseases. The combination may also influence the complete characteristics along with the data cleaning mechanism. In the case of overall automation, accuracy, and error rate the combination of data mining, machine learning, and evolutionary algorithms are found to be prominent. But the major significance in these approaches is also the combinatorial use related to every method (Figure 15). The complexity of the procedures does not affect the overall result emphasis as the result is the major motivation in case of disease prediction and detection.

The symptoms of health-related issues may vary from person to person and case by case so the detection and prediction strategy should be changed. So, the identification of the accurate computational method

is a bigger challenge. The association between symptoms and the internal and external impacts should be acquired to gather accurate information. So, the association between these aspects is also a major challenge. The suitability of each method with different problem scenarios should be tested and validated based on some dummy creation. It may help in handling real events efficiently.

The association between risk factors and mental health problems should be compared and analyzed for the early-stage prediction. Most of the mental disorders are macro issues but the risk associated may be high so there is the need for interventions with each symptom for the prevention and prediction. There is also the need for removal or addition of such symptoms which may affect the causes and the risk in positive and negative detection.

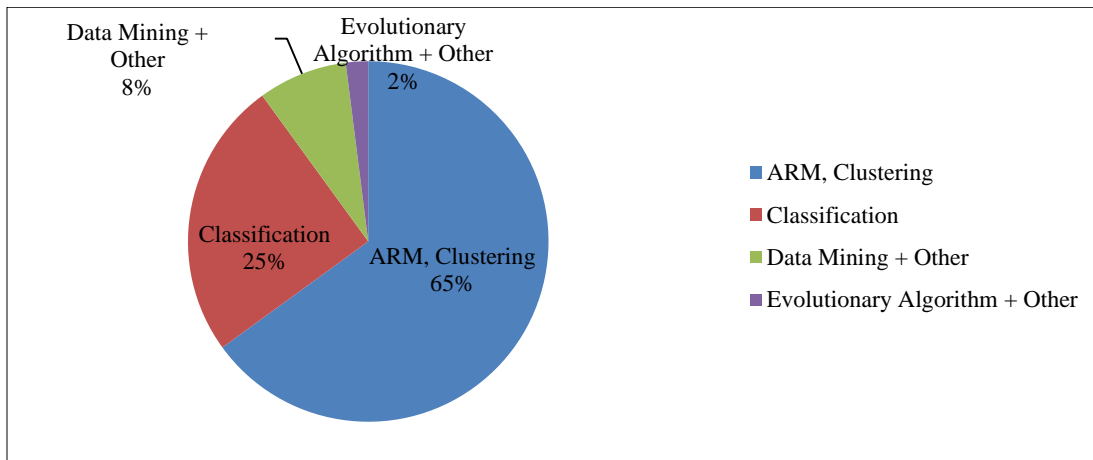


Figure 13 Methods used in pre-processing and pattern extraction for the health data prediction

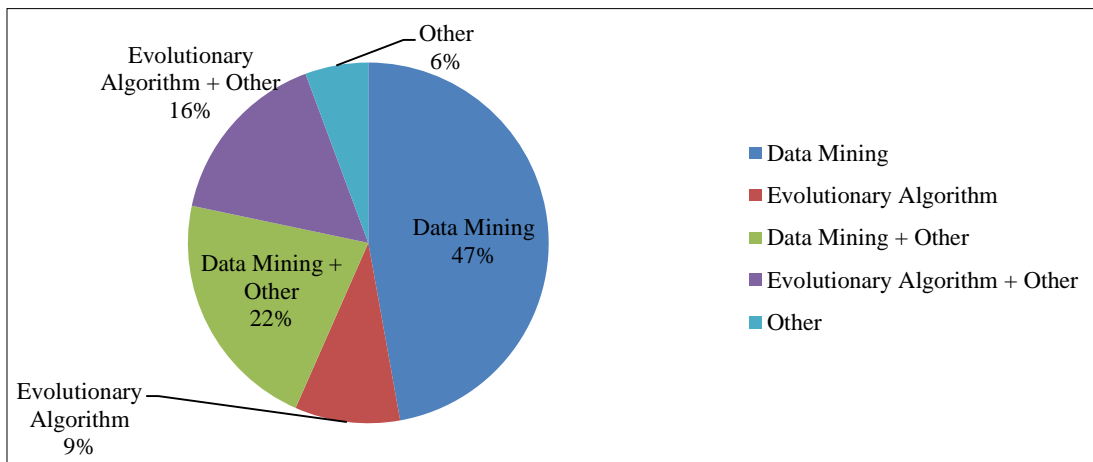


Figure 14 Methods used in final accuracy for the health data prediction

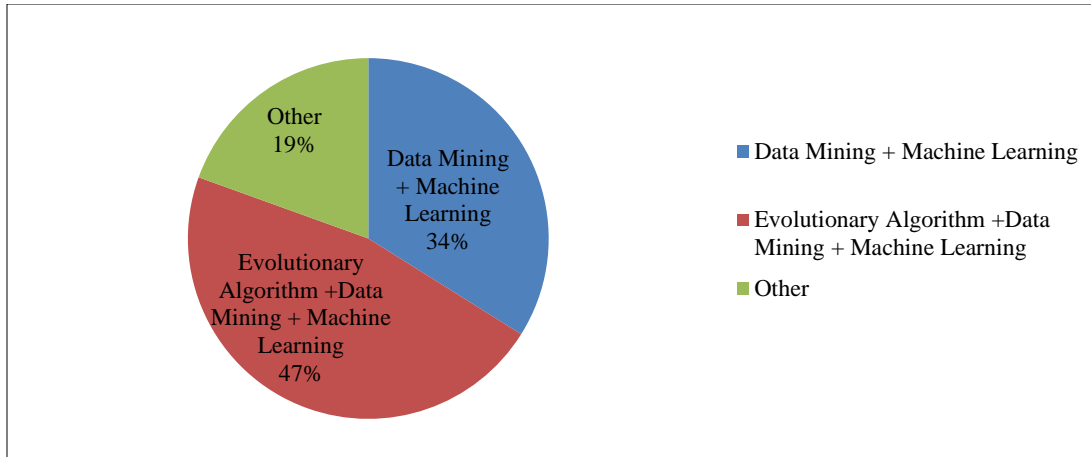


Figure 15 Complete evaluation for the automated process for the health data prediction

Limitations

The following limitations have been analyzed based on the complete study.

1. There is the need for considering symptom-based parameters which are related to the diseases or the health care system considered, as it can vary based on nature.
2. The structure of the health care system may be different, so it should be considered in terms of individual cases and solved separately [96].
3. Computational aspects should be considered with the environmental factors also as it can vary in terms of depression or other related diseases.
4. The individual and associative influences of each considered parameters should be considered for the final validation.
5. How to apply the combination and the incorporation between the label management should be considered separately along with the joint view to compare the performances [96].

5. Conclusion

In this paper DSS and ICT system for health informatics, approaches have been discussed. It also includes different Computational and methodological exploration in terms of decision making and prediction related to the health care system and disease data. This paper elaborates the latest trends in terms of computing health care considering ICT, data mining, and machine learning approaches. It elaborates the computational, analytical along with methodological prospects for the understanding of the role of DSS and ICT along with the combination of other techniques. It also includes the methodological view in terms of method detail, explanation, algorithm, and real-time applicability. This paper shows the collaboration of computational

mechanisms for the betterment of health care informatics. The approaches considered here are data mining, evolutionary algorithm, and machine learning. It has been considered as the performance parameters for disease detection and prediction in the health care system. These approaches have been discussed with the computational parameters, diseases specific relations, symptoms-based cases, and performance analysis. Based on the discussion and analysis gaps have been identified with the suggested measures and the limitations in the traditional methodology.

Acknowledgment

None.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

- [1] Zonneveld M, Patomella AH, Asaba E, Guidetti S. The use of information and communication technology in healthcare to improve participation in everyday life: a scoping review. *Disability and Rehabilitation*. 2020; 42(23):3416-23.
- [2] Fischl C, Asaba E, Nilsson I. Exploring potential in participation mediated by digital technology among older adults. *Journal of Occupational Science*. 2017; 24(3):314-26.
- [3] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*. 2009; 62(10):E1-34.
- [4] Nyasulu C, Chawinga WD. The role of information and communication technologies in the delivery of health services in rural communities: experiences from

- Malawi. South African Journal of Information Management. 2018; 20(1):1-10.
- [5] Teng X, Gong Y. Research on application of machine learning in data mining. Materials science and engineering. 2018 (pp.1-4) IOP Publishing
- [6] Ayanso A, Lertwachara K. An analytics approach to exploring the link between ICT development and affordability. Government Information Quarterly. 2015; 32(4):389-98.
- [7] Ganser M, Dhar S, Kurup U, Cunha C, Gacic A. A data-driven approach towards patient identification for telehealth programs. In international conference on big data 2015 (pp. 2551-9). IEEE.
- [8] Fuller RM, Vician CM, Brown SA. Longitudinal effects of computer-mediated communication anxiety on interaction in virtual teams. IEEE Transactions on Professional Communication. 2016; 59(3):166-85.
- [9] Cai J, Wang ZJ, Appel-Cresswell S, Mckeown MJ. Feature selection to simplify BDI for efficient depression identification. In Canadian conference on electrical and computer engineering 2016 (pp. 1-4). IEEE.
- [10] Alabdani OM, Aldahash AA, AlKhalil LY. A framework for depression dataset to build automatic diagnoses in clinically depressed Saudi patients. In SAI computing conference 2016 (pp. 1186-9). IEEE.
- [11] Bochicchio M, Cuzzocrea A, Vaira L. A big data analytics framework for supporting multidimensional mining over big healthcare data. In international conference on machine learning and applications 2016 (pp. 508-3). IEEE.
- [12] Dhaka P, Johari R. HCAB: healthcare analysis and data archival using big data tool. In international conference on information processing 2016 (pp. 1-6). IEEE.
- [13] Li X, Sedeh RS, Wang L, Yang Y. Patient-record level integration of de-identified healthcare big databases. In international conference on big data 2016 (pp. 1784-6). IEEE.
- [14] Koppad SH, Kumar A. Application of big data analytics in healthcare system to predict COPD. In international conference on circuit, power and computing technologies 2016 (pp. 1-5). IEEE.
- [15] Guo Y, Zhang H, Pang C. EEG-based mild depression detection using multi-objective particle swarm optimization. In Chinese control and decision conference 2017 (pp. 4980-4). IEEE.
- [16] Chaitanya MN, Jayakkumar S, Chong E, Yeow CH. A wearable, EEG-based massage headband for anxiety alleviation. In annual international conference of the engineering in medicine and biology society 2017 (pp. 3557-60). IEEE.
- [17] Costarides V, Kouris I, Anastasiou A, Koutsouris DD. An innovative gaming approach to prevent anxiety disorders & promote youth resilience. In international symposium on computer-based medical systems 2017 (pp. 191-2). IEEE.
- [18] Hajian M, Moradi MH. Quantification of depression disorder using EEG signal. In national and international Iranian conference on biomedical engineering. 2017 (pp. 1-5). IEEE.
- [19] Haritha H, Negi S, Menon RS, Kumar AA, Kumar CS. Automating anxiety detection using respiratory signal analysis. In region 10 symposium 2017 (pp. 1-5). IEEE.
- [20] Li LK, Li YX, Hung CY. Multi-touch technology in education: digital simulation improves student comprehension, cognitive concept, and anxiety. In international symposium on educational technology 2019 (pp. 241-4). IEEE.
- [21] Mazlan A, Bolong J, Nasrudin SR, Rahman AH. The impact of social interaction anxiety in the use of learning management system: a tentative model. In international conference on engineering technology and technopreneurship 2017 (pp. 1-4). IEEE.
- [22] Alharbi AH. Health informatics e-learning object repository HiLOR. In international conference on informatics, health & technology 2017 (pp. 1-7). IEEE.
- [23] Faruk N, Surajudeen-Bakinde NT, Oloyede AA, Bello OO, Popoola SI, Abdulkarim A, et al. On green virtual clinics: a framework for extending health care services to rural communities in Sub-Saharan Africa. In international rural and elderly health informatics conference 2017 (pp. 1-7). IEEE.
- [24] Chatterjee P, Cymberknop LJ, Armentano RL. IoT-based decision support system for intelligent healthcare—applied to cardiovascular diseases. In international conference on communication systems and network technologies 2017 (pp. 362-6). IEEE.
- [25] Ghoshachandra P, Limkriengkrai C, Wimsakcharoen P, Tangsripairoj S. oHealth: a self-care android application for senior citizens with hypertension. In ICT international student project conference 2017 (pp. 1-5). IEEE.
- [26] Islam M, Ashikuzzaman M, Tabassum T, Yusuf MS. A non-invasive technique of early heart diseases prediction from photoplethysmography signal. In international conference on electrical information and communication technology 2017 (pp. 1-5). IEEE.
- [27] Pouriyeh S, Vahid S, Sannino G, De Pietro G, Arabia H, Gutierrez J. A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease. In symposium on computers and communications 2017 (pp. 204-7). IEEE.
- [28] Priyanga P, Naveen NC. Web analytics support system for prediction of heart disease using naive bayes weighted approach (NBWA). In Asia modelling symposium 2017 (pp. 21-6). IEEE.
- [29] Yeh SC, Li YY, Zhou C, Chiu PH, Chen JW. Effects of virtual reality and augmented reality on induced anxiety. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2018; 26(7):1345-52.
- [30] Lee S, Lee T, Yang T, Seomoon E, Yoon C, Kim SP. Neural correlates of anxiety induced by environmental events during driving. TENCON 2018 (pp. 0446-50). IEEE.
- [31] Aceto G, Persico V, Pescapé A. The role of information and communication technologies in

- healthcare: taxonomies, perspectives, and challenges. *Journal of Network and Computer Applications*. 2018; 107:125-54.
- [32] Liu H, Wen W, Zhang J, Liu G, Yang Z. Autonomic nervous pattern of motion interference in real-time anxiety detection. *IEEE Access*. 2018; 6:69763-8.
- [33] Fritchman K, Saminathan K, Dowsley R, Hughes T, De Cock M, Nascimento A, et al. Privacy-preserving scoring of tree ensembles: a novel framework for AI in healthcare. In *international conference on big data 2018* (pp. 2413-22). IEEE.
- [34] Breil B, Kremer L, Taweel A, Lux T. A comparative literature analysis of the health informatics curricula. In *international conference on computer systems and applications 2018* (pp. 1-4). IEEE.
- [35] Khader Y, Jabareen H, Alzyoud S, Awad S, Rumeileh NA, Manasrah N, et al. Perception and acceptance of health informatics learning among health-related students in Jordan and Palestine. In *international conference on computer systems and applications 2018* (pp. 1-6). IEEE.
- [36] Kaur S, Bawa RK. Review on data mining techniques in healthcare sector. In *international conference on I-SMAC (IoT in social, mobile, analytics and cloud) 2018* (pp. 224-8). IEEE.
- [37] Lee TD, Park H, Lee J. Collaborative accountability for sustainable public health: a Korean perspective on the effective use of ICT-based health risk communication. *Government Information Quarterly*. 2019; 36(2):226-36.
- [38] Kimani E. A sensor-based framework for real-time detection and alleviation of public speaking anxiety. In *8th international conference on affective computing and intelligent interaction workshops and demos 2019* (pp. 55-9). IEEE.
- [39] Li X. DURS: a distributed method for k-nearest neighbor search on uncertain graphs. In *international conference on mobile data management 2019* (pp. 377-8). IEEE.
- [40] Ma L, Wang Y. Constructing a semantic graph with depression symptoms extraction from twitter. In *conference on computational intelligence in bioinformatics and computational biology 2019* (pp. 1-5). IEEE.
- [41] Amith M, Cui L, Roberts K, Xu H, Tao C. Ontology of consumer health vocabulary: providing a formal and interoperable semantic resource for linking lay language and medical terminology. In *international conference on bioinformatics and biomedicine 2019* (pp. 1177-8). IEEE.
- [42] Balatsoukas P, Porat T, Sassoon I, Essers K, Kokciyan N, Chapman M, et al. User involvement in the design of a data-driven self-management decision support tool for stroke survivors. In *EUROCON 2019-18th international conference on smart technologies 2019* (pp. 1-6). IEEE.
- [43] Yahyaoui A, Jamil A, Rasheed J, Yesiltepe M. A decision support system for diabetes prediction using machine learning and deep learning techniques. In *international informatics and software engineering conference 2019* (pp. 1-4). IEEE.
- [44] Vogel D. Adults use of ICT in healthcare: the persuasive impact of children. *Procedia Computer Science*. 2016; 98:236-42.
- [45] Stoicu-Tivadar L. ICT frame supporting continuity of care towards increased quality of healthcare services. *IFAC Proceedings Volumes*. 2010; 43(23):180-3.
- [46] Melzer A, Onic A, Starzer F, Huemer M. Short-range leakage cancelation in FMCW radar transceivers using an artificial on-chip target. *IEEE Journal of Selected Topics in Signal Processing*. 2015; 9(8):1650-60.
- [47] Zeadally S, Bello O. Harnessing the power of internet of things based connectivity to improve healthcare. *Internet of Things*. 2019.
- [48] Yadav M, Sakib MN, Feng K, Chaspari T, Behzadan A. Virtual reality interfaces and population-specific models to mitigate public speaking anxiety. In *international conference on affective computing and intelligent interaction 2019* (pp. 1-7). IEEE.
- [49] Puli A, Kushki A. Toward automatic anxiety detection in autism: a real-time algorithm for detecting physiological arousal in the presence of motion. *IEEE Transactions on Biomedical Engineering*. 2019; 67(3):646-57.
- [50] Cacheda F, Fernandez D, Novoa FJ, Carneiro V. Early detection of depression: social network analysis and random forest techniques. *Journal of Medical Internet Research*. 2019; 21(6):1-18.
- [51] Wang S, Wang J, Wang X, Qiu T, Yuan Y, Ouyang L, et al. Blockchain-powered parallel healthcare systems based on the ACP approach. *IEEE Transactions on Computational Social Systems*. 2018; 5(4):942-50.
- [52] Meephak N, Tepbanchaporn M, Jarupaibul A. Elder eat: a smartphone application for recording and monitoring food consumption for thai elderly. In *ICT international student project conference 2018* (pp. 1-6). IEEE.
- [53] Navin K, Krishnan MM, Lavanya S, Shanthini A. A mobile health based smart hybrid epidemic surveillance system to support epidemic control programme in public health informatics. In *international conference on IoT and application 2017* (pp. 1-4). IEEE.
- [54] Movahedi F, Zhang Y, Padman R, Antaki JF. Mining temporal patterns from sequential healthcare data. In *international conference on healthcare informatics 2018* (pp. 461-2). IEEE.
- [55] Malavasi M., Cesario L., Fiordelmondo V., Gherardini A., Hoogerwerf E.J., Lepore C., et al., Designing, implementing and testing an IoT based home system for integrated care services: exploring the transferability and expandibility of an ICT based integrated care system to respond to the long term daily needs of people with disabilities. In *international symposium on consumer technologies*. 2019 (pp. 53-6). IEEE.
- [56] Xia E, Wang K, Zhang Y, Yu Y, Mei J, Li S. A data-driven clinical decision support system for acute coronary syndrome patient similarity. In *international*

- conference on healthcare informatics 2019 (pp. 1-6). IEEE.
- [57] Chisengantambu-Winters C, Robinson GM, Evans N. Developing a decision-making dependency (DMD) model for nurse managers. *Heliyon*. 2020; 6(1):1-10.
- [58] Van De Walle B, Turoff M. Decision support for emergency situations. In *handbook on decision support systems 2008* (pp. 39-63). Springer, Berlin, Heidelberg.
- [59] Gourbesville P. Integrated river basin management, ICT and DSS: challenges and needs. *Physics and Chemistry of the Earth, Parts A/B/C*. 2008; 33(5):312-21.
- [60] Lindblom J, Lundström C, Ljung M, Jonsson A. Promoting sustainable intensification in precision agriculture: review of decision support systems development and strategies. *Precision Agriculture*. 2017; 18:309-31.
- [61] Yi L, Thomas HR. A decision support system for the environmental impact of ICT and e-business. *International Journal of Information Technology & Decision Making*. 2009; 8(2):361-77.
- [62] Huang D, Wang CD, Wu JS, Lai JH, Kwoh CK. Ultra-scalable spectral clustering and ensemble clustering. *IEEE Transactions on Knowledge and Data Engineering*. 2019; 32(6):1212-26.
- [63] Shridhar M, Parmar M. Survey on association rule mining and its approaches. *International Journal of Computer Sciences and Engineering*. 2017; 5(3):129-35.
- [64] Huang D, Wang CD, Lai JH. Locally weighted ensemble clustering. *IEEE Transactions on Cybernetics*. 2017; 48(5):1460-73.
- [65] Patel S, Sihmar S, Jatain A. A study of hierarchical clustering algorithms. In *international conference on computing for sustainable global development 2015* (pp. 537-41). IEEE.
- [66] Gupta T, Panda SP. Clustering validation of CLARA and K-Means using silhouette & DUNN measures on iris dataset. In *international conference on machine learning, big data, cloud and parallel computing 2019* (pp. 10-3). IEEE.
- [67] Kapil S, Chawla M, Ansari MD. On K-means data clustering algorithm with genetic algorithm. In *fourth international conference on parallel, distributed and grid computing 2016* (pp. 202-6). IEEE.
- [68] Shakiba A. Differentially private fuzzy C-means clustering algorithms for fuzzy datasets. In *iranian joint congress on fuzzy and intelligent systems*. 2018 (pp. 91-3). IEEE.
- [69] Fijri AL, Rustam Z. Comparison between fuzzy kernel C-means and sparse learning fuzzy C-means for breast cancer clustering. In *international conference on applied information technology and innovation 2018* (pp. 158-61). IEEE.
- [70] Sinaga KP, Hsieh JN, Benjamin JB, Yang MS. Modified relational mountain clustering method. In *international conference on artificial intelligence and soft computing 2018* (pp. 690-701). Springer, Cham.
- [71] Ngo LT, Pham BH. A type-2 fuzzy subtractive clustering algorithm. In *mechanical engineering and technology 2012* (pp. 395-402). Springer, Berlin, Heidelberg.
- [72] Sharma A, Sharma MK, Dwivedi RK. Analyzing facebook data set using self-organizing map. In *international conference on system modeling & advancement in research trends 2018*. (pp. 109-12). IEEE
- [73] Pillai NS, Ranihemamalini R, Agila G, Pavithra N. Satellite image classification using self organizing map and ensemble classifiers. In *international conference on system, computation, automation and networking 2019* (pp. 1-5). IEEE.
- [74] Dubey AK, Gupta U, Jain S. Analysis of k-means clustering approach on the breast cancer Wisconsin dataset. *International Journal of Computer Assisted Radiology and Surgery*. 2016; 11(11):2033-47.
- [75] Dubey AK, Gupta U, Jain S. Comparative study of K-means and fuzzy C-means algorithms on the breast cancer data. *International Journal on Advanced Science, Engineering and Information Technology*. 2018; 8(1):18-29.
- [76] Dubey AK, Kumar A, Agrawal R. An efficient ACO-PSO-based framework for data classification and preprocessing in big data. *Evolutionary Intelligence*. 2020.
- [77] Chahar R, Kaur D. A systematic review of the machine learning algorithms for the computational analysis in different domains. *International Journal of Advanced Technology and Engineering Exploration*. 2020; 7(71): 147-64.
- [78] Hussin SK, Omar YM, Abdelmageid SM, Marie MI. Traditional machine learning and big data analytics in virtual screening: a comparative study. *International Journal of Advanced Computer Research*. 2020; 10(47):72-88.
- [79] Kumari I, Sharma V. A review for the efficient clustering based on distance and the calculation of centroid. *International Journal of Advanced Technology and Engineering Exploration*. 2020; 7(63):48-52.
- [80] Wu Y. A new instance-weighting naive bayes text classifiers. In *international conference of intelligent robotic and control engineering 2018* (pp. 198-202). IEEE.
- [81] Wu L, Li M. Applying the CG-logistic regression method to predict the customer churn problem. In *international conference on industrial economics system and industrial security engineering 2018* (pp. 1-5). IEEE.
- [82] Feng D, Deng Z, Wang T, Liu Y, Xu L. Identification of disturbance sources based on random forest model. In *international conference on power system technology 2018* (pp. 3370-5). IEEE.
- [83] Huang Y, Capretz LF, Ho D. Neural network models for stock selection based on fundamental analysis. In *Canadian conference of electrical and computer engineering 2019* (pp. 1-4). IEEE.

- [84] Ibrahim D. An overview of soft computing. *Procedia Computer Science*. 2016; 102:34-8.
- [85] Saerens M, Soquet A. Neural controller based on back-propagation algorithm. In *IEE proceedings F (Radar and Signal Processing)* 1991; 138(1):55-62. IET Digital Library.
- [86] Bae C, Wahid N, Chung YY, Yeh WC, Bergmann NW, Chen Z. Effective audio classification algorithm swarm-based optimization. *International Journal of Innovative Computing, Information and Control*. 2014; 10(1):151-67.
- [87] Rao RV, Savsani VJ, Vakharia DP. Teaching-learning-based optimization: a novel method for constrained mechanical design optimization problems. *Computer-Aided Design*. 2011; 43(3):303-15.
- [88] Kennedy J, Eberhart R. Particle swarm optimization. In *proceedings of ICNN'95-international conference on neural networks 1995* (pp.1942-8). IEEE.
- [89] Shi Y, Eberhart R. A modified particle swarm optimizer. In *international conference on evolutionary computation proceedings 1998* (pp. 69-73). IEEE.
- [90] Dubey AK, Narang S, Kumar A, Sasubilli SM, García-Díaz V. Performance estimation of machine learning algorithms in the factor analysis of covid-19 dataset. *Computers, Materials & Continua*. 2021; 66(2):1921–36.
- [91] Dorigo M. Optimization, learning and natural algorithms. PhD Thesis, Politecnico di Milano. 1992.
- [92] Dorigo M, Gambardella LM. Ant colony system: a cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*. 1997; 1(1):53-66.
- [93] Rao RV, Patel V. An improved teaching-learning-based optimization algorithm for solving unconstrained optimization problems. *Scientia Iranica*. 2013; 20(3):710-20.
- [94] Dileep G, Singh SN. Application of soft computing techniques for maximum power point tracking of SPV system. *Solar Energy*. 2017; 141:182-202.
- [95] Falcone R, Lima C, Martinelli E. Soft computing techniques in structural and earthquake engineering: a literature review. *Engineering Structures*. 2020.
- [96] Dubey AK, Gupta U, Jain S. Computational measure of cancer using data mining and optimization. In *international conference on sustainable communication networks and application 2019* (pp. 626-32). Springer, Cham.



Ravita Chahar is an Assistant Professor in Computer Science and Engineering Department at Chitkara University, Punjab. She is having a rich experience of 12 years in the teaching field. She holds the degree of Master in Engineering in Computer Science and Engineering from Maharshi Dayanand University, Rohtak. Her area of research interests are Databases, Data Mining and Machine Learning. Currently, she is pursuing Ph.D in Computer Science and Engineering from the Chitkara University, Punjab.
Email: ravita.chahar@chitkara.edu.in