Review on Clinical Decision Support System for Electronic Health Record System for Major Diseases

Punam S. Pawar
ME Student of CE and IT department
R. C. Patel Institute of Technology
Shirpur, India
punamspawar@gmail.com

D. R. Patil
Faculty of CE and IT department
R. C. Patel Institute of Technology
Shirpur, India
dharmaraj.rcpit@gmail.com

***Abstract*—Patient is the main source of information to physicians while treating them. Professionals diagnose patients depending on their experience, guidelines. But when they are treating a patient who is having major disease or if the patient is serious due to accident i.e. in emergency situation then proper decision making is essential one which will provide correct diagnosis, medications. And there are more chances to take wrong decision due to less time and less data related to patient. In that case clinical decision support system and electronic health record systems can perform major role for taking appropriate decisions as correct data is available in less time in place.**

Keywords-electronic health record (EHR); decision support system (DSS); clinical decision support system (CDSS); neural network.

I. INTRODUCTION

The World Health Organization (WHO) defines Non-communicable Diseases (NCDs) as including chronic diseases (principally cardiovascular disease, diabetes, cancer, and asthma/chronic respiratory disease), injuries, and mental health. One of the leading causes of death and disability in India today: non-communicable diseases like heart attacks and strokes, cancers, diabetes and chronic respiratory disease. Cardiovascular diseases including stroke, diabetes, cancer, and chronic respiratory disease account for 80% of all deaths due to NCDs. India is having most people with diabetes, with a current figure of 50.8 million, followed by China with 43.2 million [1]. Type 1 diabetes is usually diagnosed in children or young adults, although it can occur at any age. Approximately 5-10 % of all people with diabetes are diagnosed with Type 1. Type 2 diabetes is most common in people older than 45 who are overweight. However, as a consequence of increased obesity among the young, it is becoming more common in children and young adults. Type 2 diabetes is the most common type of diabetes and accounts for 90-95% of all diabetes. Some women develop a third, usually temporary, type of diabetes called 'gestational diabetes' when they are pregnant. Gestational diabetes develops in 2-5% of all pregnancies but usually disappears when the pregnancy is over. Women who have had gestational diabetes have an increased risk of developing type 2 diabetes later on. According to the Causes of Death Survey (2001- 03) conducted by the Registrar General of India, non-communicable diseases are the leading causes of death in the country, constituting 42% of all deaths (urban 56%; rural 40%). The estimated burden of various NCDs in the country is tabulated below in Table 1. These are estimates based on the best available information.

TABLE I
ESTIMATED BURDEN OF VARIOUS NCDs IN COUNTRY

Disease	Estimated cases	Estimamted Annual Mortality	Source
Cardiovascular disorders	27 million (2000) 61 million (2015)	2.7 million	National Commission on Macroeconomics and Health
Diabetes	51 million (2010)	0.1 million	IDF Diabetes Atlas, 4th Edition
Stroke	1.5 million	0.6 million	National Commission on Macroeconomics and Health
Chronic Repository Disorders	>30 million (2010)	0.8 million	INSEARCH report to ICMR
Cancers	0.95 million incident cases	0.6 million	GLOBOCAN 2008

II. DECISION SUPPORT SYSTEM

A DSS can be described as a computer-based interactive human–computer decision-making system that:

A. *supports decision makers rather than replaces them;*

B. *utilizes data and models;*

C. *solves problems with varying degrees of structure:*
   1) non-structured (unstructured or ill-structured) (Bonczek *et al.* 1981);
   2) semi-structured (Keen and Scott-Morton 1978);
   3) semi-structured and unstructured (Sprague and Carlson 1982);

D. focuses on effectiveness rather than efficiency in decision processes (facilitating decision processes).

According to Sprague and Carlson [2], decision support systems would consist in the following components (Fig. 1): data management component; model management component; user interface management component; decision support system architecture.

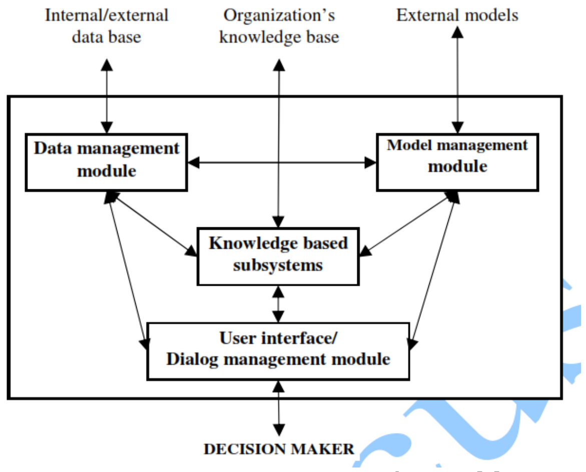

Figure 1. Decision support system architecture [2]

Power's differentiation is used [3] which divides DSS's into five groups:
•Communication-driven DSS – helps in a group task by supporting communication among workers;
•Data-driven DSS – concentrates on the access and manipulation of both the internal and external data;
•Document-driven DSS – applies to unstructured information that is managed, retrieved and manipulated with the use of the system into a variety of electronic formats;
• Knowledge-driven DSS – specialized in solving problems basing on facts, rules or similar constructions;
•Model-driven DSS – puts emphasis on simulation, financial support and optimization tools that are based on statistical solutions.

III. CLINICAL DECISION SUPPORT SYSTEM

It is also called as medical decision support systems (MDSSs). Clinical decision support systems (CDSSs) are computer-based information system which is useful during patient care for taking decision correctly and fast. By using EHR with CDSS the patient safety can be improved.

A. Definitions of CDSSs:

In the literature, many researchers have given their definitions of CDSSs. Some typical definitions are given below. Musen[7] defined a CDSS as any piece of software that takes information about a clinical situation as inputs and that produces inferences as outputs that can assist practitioners in their decision making and that would be judged as "intelligent" by the program's users. Miller and Geissbuhler[8] defined a CDSS providing diagnostic decision support as a computer-based algorithm that assists a clinician with one or more component steps of the diagnostic process. Sim et al.[9] defined CDSSs as "software that designed to be a direct aid to clinical decision-making, in which the characteristics of an individual patient are matched to a computerized clinical knowledge base and patient specific assessments or recommendations are then presented to the clinician or the patient for a decision".

B. A General Model of CDSS:

A general model of CDSSs which has been discussed in the literature is shown as Fig. 2.

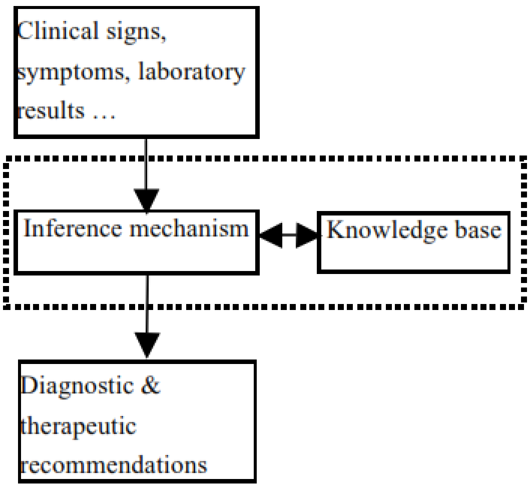

Figure 2. A General Model of CDSS

The model has clinical signs, symptoms, and laboratory results as inputs and output as diagnostic and therapeutic recommendations. For that it uses inference mechanism and knowledge base.

C. Types of CDSS:

According to Metzger and her colleagues have described CDSSs using several dimensions ([4], [5]):

- According to their framework, CDSSs differ among themselves in the timing at which they provide support (before, during or after that clinical decision is made) and

- How active or passive the support is, i.e. whether CDSS actively provides alerts or passively responds to physician input or patient-specific information.

- CDSS vary in how easy they are for busy clinicians to access.

- CDSS also differ in whether the information provided is general or specialty-based.

- Another categorization scheme for CDSSs is whether they are knowledge-based systems or non-knowledge-based systems that employ machine learning or other statistical pattern recognition approaches.

IV. LITERATURE SURVEY

Problem Oriented Medical Records (POMR) and Subjective Objective Assessment Plan (SOAP) methods were proposed by Lawrence Weed.

Liang Xiao et al. [10] used problem oriented approach due to two major advantages that are extensibility and decision support. They design EHR for clinical decision support by hierarchical EHR model having three levels with international standards ICPC 2 (International Classification of Primary Care-Second edition), openEHR archetype and SNOMED CT (Systematized Nomenclature of Medicine-Clinical Terms) for the methadone treatment protocol.

In the first level authors use paper-based “Addiction Assessment Form” guideline for the construction of baseline archetype. In the evaluation step General Practitioners establishes knowledge about patient’s heroin habit and forms current management plans. The next step i.e. prescribing instruction is constructed from paper-based “Methadone Prescription Form”. Depending on that prescription instruction appropriate action i.e. the last step is performed.

In the last level 3 the archetypes are link to terminology using standard SNOMED CT.

Authors used web-based data entry system for implementation. For the ease of implementation authors represent the knowledge of guidelines as decision making flowchart. As per their opinion their system is more useful to less experience doctors who need frequent reference to guidelines.

Zeynab Khalifelu et al. [11] used neural network for diagnosis of patient specifically for heart disease. The architecture used is the multilayer perceptron network and the back-propagation algorithm is used as learning algorithm. They finalised the number of input nodes, number of hidden nodes and number of output nodes with the help of paper [6].

Authors assumed 2 classes for their neural network model that are Positive class for the existence of heart failure and negative class means heart failure not exist. Their results indicate that the system with neural network for 85% test sets predicate correct answers.

Mrudula Gudadhe et al. [12] given a DSS for the identification of heart disease by using support vector machine and artificial neural network.

Authors use Cleveland Heart Database taken from UCI learning data set repository which was donated by Detrano [13], which has 13 numeric attributes. For the training of multilayer perceptron the back propagation algorithm is used. As per their results both methods have high accuracy but support vector machine classifies heart disease data into two classes i.e. only presence or absence of disease, while artificial neural network classifies that data into 5 categories of heart failure with more accuracy than support vector machine.

Sri Hartati [14] gives Kohonen artificial neural network as a DSS model for prediction of coronary artery disease (CAD.

Authors use Kohonen Self-organizing map (KSON) to predict whether patient suffer from coronary artery disease or not based on risk factors which belongs to patient. The KSON update its weight features without need for performance feedback from teacher or network trainer. The feature of KSON is that the nodes distribute themselves across input space to identify groups of similar input vectors which in this case values representing the risk factors of CAD. The learning process is based on competitive learning techniques i.e. winner take all strategy. This algorithm when trained, produce a low dimension representation of input space. It uses steps computation in training algorithm. For that it chooses some input pattern from input data set. The index of winning unit is given by Euclidian distance is minimized. Weights are updated according to the global network updating phase from iteration to iteration. The learning rate and neighbourhood are decreased at each iteration by using shrinking function.

These iterations are continued until each output reach upto threshold of sensitivity w.r.t. input portion space. Input vectors are grouped into classes by calculating distance between optimal weight to input vector. This model can be used by doctors and patients in CAD diagnosis.

Ping-Tsai Chung et al. [15] gives intelligent diagnosis and also gives intelligent diagnosis and also gives inferential advices for interrelated disease using a knowledge based DSS for healthcare. Fig. 3 gives architecture of knowledge-based decision system.

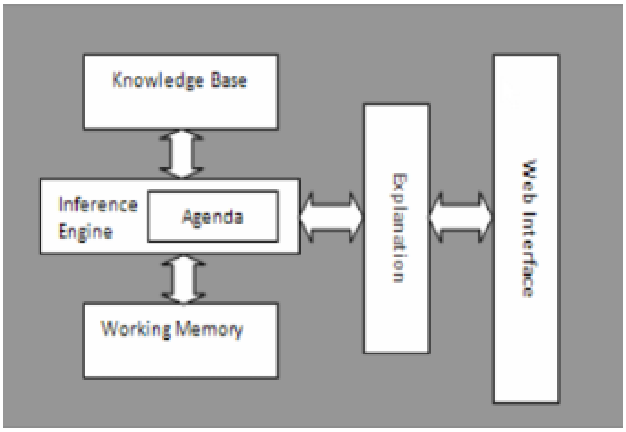

Figure 3. Architecture of Knowledge-based DSS [15]

The knowledge base is the heart of it. Their system provides comprehensive Discoveries and Advices by using Chaining Inferential Technique. By providing healthy advice to people it promotes quality of life. The system presented by them is portable and compatible with most of major platforms. For knowledge acquisition the process of extraction, organisation and structuring knowledge is performed. After it authors use Rules for knowledge representation which could be understandable to computer.

Their system provides diagnosis and advice for Body Mass Index, Diabetes, Blood Pressure, and Cholesterol. It also gives deep diagnosis 1 for unknown type of diabetes and deep diagnosis 2 for Asian Pregnant Woman. Authors have conducted survey for healthcare knowledge based DS. They extend their knowledge base of intelligence decision system for 10-years risk prediction.

Mohd Fauzi Othman et al. [16] have given probabilistic neural network (PNN) for brain tumour classification. Authors used Principal Component Analysis (PCA) technique for feature extraction. This technique is use for the reduction of dimensions of data. In the training phase the PCA is used to reduce the dimensionality of feature vector which are extracted for each image, which is in training set.

While in testing phase, the feature vector of test image computed using PCA is composed with all feature vectors of training sets which are computed previously. The similarities are computed with Euclidian distance. Due to the easy training and instanteoussness authors used the probabilistic neural network. Authors have used 3 layers of PNN are: input layer, radial basis layer and competitive layer.

Authors used PNN for classification of MR image data. Their experiment results indicate that PNN gives accuracy from 100% to 73% depending on spread value.

V. METHODOLOGY

In the proposed decision support system for the classification of cancer disease we will use multilayer perceptron architecture of artificial neural network. We will use back propagation network as it is a fast learning algorithm for weight adjustment.

A. Multi Layer Perceptron:

The single layer perceptron is not able to solve non-linearly separable problems. For that purpose one or more layers are added in single layer perceptron i.e. multi-layer perceptron. Multilayer perceptron network is a feed-forward neural network as shown in Fig. 4. They are widely used for pattern classification, recognition, prediction and approximation.

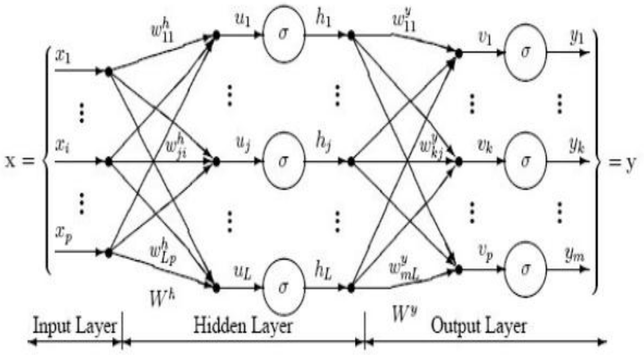

Figure 4. Multi-layer Perceptron Network [17]

Above network have an **input layer** (on the left) with three neurons, one **hidden layer** (in the middle) with three neurons and an **output layer** (on the right) with three neurons.

There is one neuron in the input layer for each predictor variable. In the case of categorical variables, *N*-1 neurons are used to represent the *N* categories of the variable.

Input Layer — a vector of predictor variable values ($x_1...x_p$) is presented to the input layer. The input layer (or processing before the input layer) standardizes these values so that the range of each variable is -1 to 1. The input layer distributes the values to each of the neurons in the hidden layer. In addition to the predictor variables, there is a constant input of 1.0, called the *bias* that is fed to each of the hidden layers; the bias is multiplied by a weight and added to the sum going into the neuron.

Hidden Layer — arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight (w_{ji}), and the resulting weighted values are added together producing a combined value u_j. The weighted sum (u_j) is fed into a transfer function, σ, which outputs a value h_j. The outputs from the hidden layer are distributed to the output layer.

Output Layer — Arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight (w_{kj}), and the resulting weighted values are added together producing a combined value v_j. The weighted sum (v_j) is fed into a transfer function, σ, which outputs a value y_k. The *y* values are the outputs of the network.

B. Back-Propagation Network:

The back-propagation is the training or learning algorithm than a network itself. The back propagation learning algorithm uses the delta-rule that is it computes the deltas, (local gradients) of each neuron starting from the output neurons and going backwards until it reaches the input layer. To compute the deltas of the output neurons, get the error of each output neuron. Since the multi-layer perceptron is a supervised training network so the error is the difference between the network's output and the desired output.

A Back Propagation network learns by example. We give to the algorithm examples of what you want the network to do and it changes the network's weights so that, when training is finished, it will give you the required output for a particular input. Back Propagation networks are ideal for simple Pattern Recognition and Mapping Tasks.

C. Learning Algorithm:

The basic MLP learning algorithm using back-propagation learning is outlined below [18]:

1. Initialise the network, with all weights set to random numbers between -1 and +1.

2. Present the first training pattern, and obtain the output.
3. Compare the network output with the target output.
4. Propagate the error backwards.

(a) Correct the output layer of weights using the following formula.

$$w_{ho} = w_{ho} + (\eta \delta_o o_h) \quad (1)$$

Where w_{ho} is the weight connecting hidden unit h with output unit o, η is the learning rate, o_h is the output at hidden unit h. δ_o is given by the following.

$$\delta_o = o_o(1 - o_o)(t_o - o_o) \quad (2)$$

where o_o is the output at node o of the output layer, and $t - o$ is the target output for that node.

(b) Correct the input weights using the following formula.

$$w_{ih} = w_{ih} + (\eta \delta_h o_i) \quad (3)$$

where w_{ih} is the weight connecting node i of the input layer with node h of the hidden layer, o_i is the input at node i of the input layer, η is the learning rate. δ_h is calculated as follows.

$$\delta_h = o_h(1 - o_h) \sum_o (\delta_h w_{ho}) \quad (4)$$

5. Calculate the error, by taking the average difference between the target and the output vector.

For example the following function could be used.

$$E = \frac{\sqrt{\sum_{n=1}^{p} (t_o - o_o)^2}}{p} \quad (5)$$

Where p is the number of units in the output layer.

6. Repeat from 2 for each pattern in the training set to complete one *epoch*.
7. Shuffle the training set randomly. This is important so as to prevent the network being influenced by the order of the data.
8. repeat from step 2 for a set number of epochs, or until the error ceases to change.

VI. CONCLUSION

The research shows that CDSSs are useful tool for physician as well as patient also because it provides right information to right person at right time in place. The main characteristics of neural networks are that they have the ability to learn complex nonlinear input-output relationships, use sequential training procedures, and adapt themselves to the data. So that it is choice of most of the researchers in CDSSs.

REFERENCES

[1] International Diabetes Federation, Press Release, Oct 19, 2009.

[2] Lungu, I, colectiv, "Sisteme informatice – Analiza, proiectare si implementare" Editura Economica, 2003, Bucuresti.

[3] Turban, E., Aronson, J.E., Liang, T.P. Decision Support Systems and intelligent Systems. New Jersey, Pearson Education, 2005.

[4] Perreault LE, Metzger JB.,A pragmatic framework for understanding clinical decision support. In: Middleton, B, ed. Clinical Decision Support Systems. J Healthc Inf Manage 1999; 13:5–21.

[5] Metzger J, MacDonald K. Clinical decision support for the independent physician practice. Oakland: California Healthcare Foundation; 2002.

[6] N. Elfadil and A. Hossen, "Identification of Patients With Congestive Heart Failure Using Different Neural Networks Approaches", Journal Technology and Health Core, vol. 17 Issue 4, December 2009.

[7] M. A. Musen, in *Handbook of medical informatics* J. H. V. a. M. Bemmel, M. A. , Ed. (Bohn Stafleu Van Loghum, Houten, 1997).

[8] R. A. Miller, A. Geissbuhler, in *Clinical Decision Support Systems* E. S. Berner, Ed. (Springer-Verlag, New York, 1999), vol. 3-34.

[9] I. Sim *et al.*, *J Am Med Inform Assoc* **8**, 527 (November 1, 2001, 2001).

[10] Liang Xiao, Grainne Cousins, Lucy Hederman, Tom Fahey, Borislav Dimitrov, "The Design of an EHR for Clinical Decision Suppport", 3rd International Conference on Biomedical Engineering and Informatics (BMEI 2010) ,2525-2531.

[11] Farhad Soleimanian Gharehchopogh, Zeynab Abbasi Khalifelu, "Neural Network Application in Diagnosis of Patient: A Case Study", IEEE,2011, 245-249.

[12] Mrudula Gudadhe, Kapil Wankhade, Snehlata Dongre, "Decision Support System for Heart Disease based on Support Vector Machine and Artificial Neural Network", International Conference on Computer and Communication Technology, 2010, 741-745.

[13] UCI Machine Learning Repository: Heart Disease Data Set. http://archive.ics.uci.edu/ml/datasets/Heart+Disease.

[14] Sri Hartati, "A Kohonen Artificial Neural Network as a DSS model for predicting CAD", International Conference on Distributed Frameworks for Multimedia Applications (DFmA), 2010.

[15] Ping-Tsai Chung, Bing-Xing Chen, "A Knowledge-Based Decision System for Healthcare Diagnosis and Advisory", IEEE, 2011, 2357-2362.

[16] Mohd Fauzi Othman, Mohd Ariffanan Mohd Basri, "Probablistic Neural Network for Brain Tumor Classification", Second International Conference on Intelligent Systems, Modelling and Simulation, 2011, 136-138.

[17] http://www.dtreg.com/mlfn.htm

[18] Leonardo Noriega, "Multilayer Perceptron Tutorial", Nov. 17, 2005.