CHAPTER 10

Neonatal Resuscitation Management using Case-Based Reasoning[†]

10.1. Introduction

The Neonates (0-4 weeks of age) are the high risk group of pediatric problem domain. The resuscitation management is most important in neonatal care. The neonatal resuscitation prevents the morbidity and mortality associated with hypoxic-ischemic tissue (brain, heart, kidney) injury and re-establishes adequate spontaneous respiration and cardiac output [1]. Approximately 5% to 10% of the newly born population require some degree of active resuscitation at birth (e.g., stimulation to breathe) [2], and approximately 1% to 10% born in the hospital are reported to require extensive ventilation [3]. More than 5 million neonatal deaths occur worldwide each year. It has been estimated that birth asphyxia accounts for 19% of these deaths, suggesting that the outcome might be improved for more than 1 million infants per year through implementation of simple resuscitative techniques [4]. Although the need for resuscitation of the newly born infant often can be predicted, such circumstances may arise suddenly and may occur in facilities that do not routinely provide neonatal intensive care. Thus, it is essential that the knowledge and skills required for resuscitation be taught to all providers of neonatal care.

Resuscitation need to be assessed for the need to receive one or more of the following actions in sequence:

A. Initial steps in stabilization (clearing the airway, positioning, stimulating)
B. Ventilation
C. Chest compressions
D. Medications or volume expansion

Progression to the next step is based on simultaneous assessment of 3 vital signs: respirations, heart rate, and colour. Progression occurs only after successful completion of the preceding step. Approximately 30 seconds is

---

[†] This is based on the publication [Journal of Association for the Advancement of Modeling and Simulation Techniques in Enterprises (AMSE), France. Advances in Modeling-C Vol.68, No.3, 2007] of the author.

allotted to complete one step successfully, re-evaluate, and decide whether to progress to the next.

Different levels of resuscitation are required depending on the signs and symptoms of a newborn as observed by a medical practitioner. In general, as per APGAR-scheme [1] of resuscitation management, these signs and symptoms lead to three classified types of illness of the newborns. Different levels of resuscitation (Level I, II and III) are required for different types of illness and the adequate treatment planning is governed.

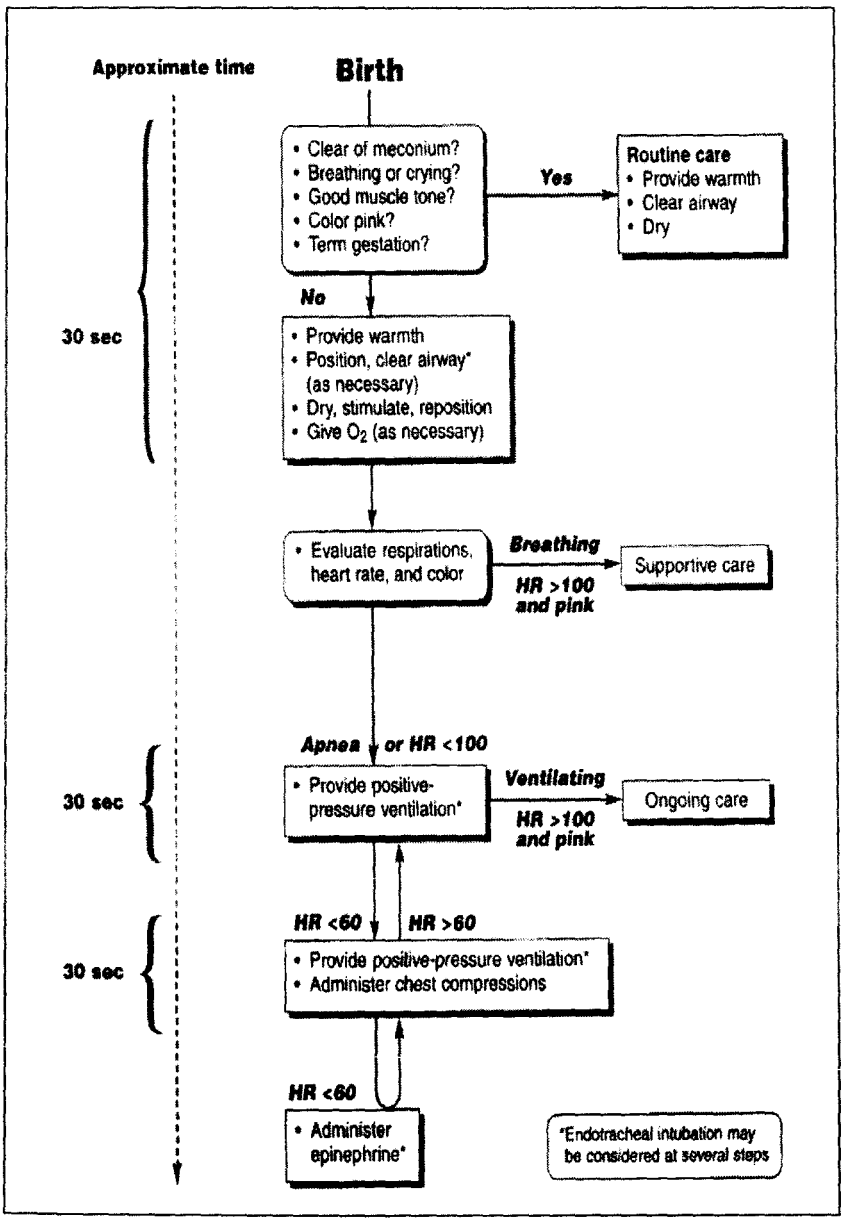

Fig. 10.1. Algorithm for resuscitation of the newly born infant [5].

For the efficient management of neonatal problems, various automated computer aided consultation system in an expert system (ES) framework have been proposed. Though rule-based expert systems are more common, but they are criticized for some shortcomings. Rules may be difficult to be elicited and formalized, unmanageably large in number, and may not be sufficient to characterize expert performance.

As an alternative, Case Based Reasoning (CBR) systems reason using analogy concepts rather than the pure decision tree (or If-THEN rules), usually adopted in rule-based systems. CBR deals with the repository of cases, and each experience and solution or outcome is retained in the case base for reuse and future reference. In medical domain, case bases may become an invaluable knowledge resource that might be utilized as a reference in diagnostic, treatment planning and outcome evaluation process. CBR methodology mirrors the type of reasoning from experience employed by medical professionals and corresponds to medical problem solving in many ways. Several medical CBR systems have already been reported in various literatures [6-12], but, however, to our knowledge, CBR has not been applied to the current problem domain of Neonatal Resuscitation Management.

Designing of a feature based CBR for diagnostic task is a two step process: proper selection of a set of potentially optimal feature subsets and design and implementation of a robust classifier. Feature selection is a preprocessing step to design a classifier. So it is encouraged to develop sophisticated feature selection methods to reduce the cost and complexity of the classifier and to improve accuracy, visualization and comprehensibility of the induced concepts [13].

Feature selection can be defined as finding a subset of relevant features among the set of original features. It has proven to be an effective means when dealing with large dimensionality with many irrelevant features [14, 15]. Feature selection in CBR for classification is well-researched problem and a lot of researches have been conducted for feature subset selection [16 - 28]. But they suffer from any one of the shortcomings: computational complexity, classifier dependency, large storage capacity and time consuming similarity measures. Moreover, most of them are concerned with the retrieval of previous cases/ instances which is a complex task.

As an alternative, another group of techniques are feature weighting approaches [29-32]. It is well known technique used to find a set of weights which assign a degree of relevance to every feature for a given task. Feature weight gives a measure of relevance. A feature is more relevant the more large the assigned weight, the irrelevant features have zero weights. The feature weighting methods determine the relevance ranking for every feature based on the rationale that a feature is likely to have different values for different classes.

Selection of features based on relevance (weight) is more appropriate for medical domain, where most of the features are presented in linguistic terms instead of numerical values. In this regard, a new scalable probabilistic filter approach proposed by Ghosh and Samanta [33] (henceforth GS model), can be considered as promising one. GS model avoids the need to select relevant features 'by hand' and gives more detail information about feature relevance by assigning weights to each context specific feature. It also suggests a limit of the number of retained cases in the case memory.

In recent years, the classification task has been enjoying increasing interest in the machine learning community. In this context, several approaches like Artificial Neural Networks, Probabilistic Neural Networks, k-Nearest Neighbor approach, Bayesian Classifier, RBF Networks etc. have been reported [34-44]. But the task of determining relevant features to the classification task is one of the central problems [45].

In our system, the classifier is designed as a single layer neural network. The task of learning is done by a CBR method and the classifier accepts the prior trained data as inputs. This model can be used for selection of features and their relevance related to the context of resuscitation management. The system becomes rich with definite known cases obtained from real field and sorts out the relevant features for all possible categories and fixes the primary weights (Relative Feature Weights) of all features through an evaluation process as suggested by Ghosh and Samanta [33]. The features are considered as input values and the Relative Feature Weights as the synapse values to the classifier respectively. The bias values are also obtained from Relative Feature Weights. When a subset of observed features is received, it is then transformed to a binary feature vector with elements 1 or 0. The input values thus processed reach to the classifier, it fires only when the weighted sum of all the potential values exceeds a certain limiting value (bias) of that particular case. The successful (fired) cases, if retained in the knowledge base, update the RFWs (synapses) of the concerned features.

The aim of this work is to apply GS model with some modification to find out the subset of features and their corresponding relevance (weight) and applying these prior trained data as input to a simple neural network based classifier to detect a particular type of illness based on the set of observations.

The next section presents an overview of the GS model which finds out the subset of features and their relevance (Relative Feature Weights) for a particular category. Section 10.3 presents the analysis of the neonatal resuscitation and architecture for feature selection. In section 10.4, the reasoning aspect is described and in 10.5, we present the theoretical foundations of the classifier. Section 10.6 presents the classifier system architecture. Section 10.7 contains case illustrations for validation of the model. We conclude in section 10.8.

10.2. An overview of GS model

The GS model is based on a case library which contains the real field cases. The case library uses a category-exemplar type [46] of case representation where knowledge is represented as cases composed of various concerned features. The same type of cases forms a category. The cases of same category are grouped in a table, called case table. To develop the system in a numerical frame, it is assumed that the features, cases and case tables are associated with some weights which represent the richness of the parameters and any decision regarding training is based on those weights. When a new case appears, all cases in the corresponding case table are scanned and an evaluation process estimates the values of some numerical parameters (weights) associated with features, cases and case tables for further judgment.

10.2.1. Definitions

When a new case appears, the initial weight of an element of feature set is not known. So an unbiased weight should be considered. This unbiased weight can be estimated by 1/N, where N is the number of elements in the set of features or sub-set of features constructing the new case. The feature weight for m-th case should be the product of average feature weight (FWav) and the number of occurrences (I) of the feature up to m-th retained case in a case table.

10.2.1.1. Feature weight (FW)

The Feature Weight (FWm) for m-th case :

$$FWm = FW(av)\ m\ .\ I \qquad (10.1)$$

$$\text{Where } FW(av)\ m = \sum_{j=1}^{M} FWj \Big/ (N.M)$$

and M is the total number of retained cases.

10.2.1.2. Case weight (CW)

It is the sum of the weights of all the possible features constructing the case. For m-th case, the Case Weight is

$$CWm = \sum Fwmi; \quad 1 \leq i \leq N \qquad (10.2)$$

where i represents the number of features contributing the m-th case.

10.2.1.3. Relative feature weight (RFW)

It is the ratio of the feature weight to the case weight. For m-th case, the relative feature weight is

$$RFWm = FWm\ /CWm \qquad (10.3)$$

10.3. System analysis and architecture for feature selection

For resuscitation management, different types of illness are identified based on the signs and symptoms of a newborn as observed by a medical practitioner. All possible sign and symptoms used for resuscitation management constitute the general feature set of the domain. Each feature possesses some defining parameters like name, type, value, weight etc as per the requirement of the problem domain.

Each case is represented with a set of features accepted as case descriptors. These sign and symptoms (features) lead to identification of 3 possible types of illness (Type-1, Type-2 and Type-3). So the case library has been designed to contain 3 categories. Each category is allotted to contain the cases of a

particular type of illness. The assigned category against each type of illness is presented in table 10.1.

Table 10.1. Allotted categories against type of illness.

Type of illness	Category
Type-1 illness	Category-1
Type-2 illness	Category-2
Type-3 illness	Category-3

The general feature set contains 34 possible sign and symptoms leading to cover these 3 types of illness mentioned above as suggested by APGAR-scheme [1] and Goswami [47]. The features are grouped under 5 rubrics as shown in table 10.2. Coding of features is used for ease of implementation.

Table 10.2. General set of features.

Serial number	Code	Features
I. Muscle Tone of Limbs		
1.	A1	Flaccid
2.	A2	Flexion at one join and rest flaccid (Hip)
3.	A3	Flexion at two joins and rest flaccid (Hip + Shoulder)
4.	A4	Flexion at three joins and rest flaccid (Hip + Shoulder + Knee/Elbow)
5.	A5	Flexion at four joins and rest flaccid (Hip + Shoulder + Knee + Elbow)
6.	A6	Flexion at five joins but no active movement (Hip + Shoulder + Knee + Elbow + Ankle/Hip)
7.	A7	Flexion at six joins but no active movement (Hip + Shoulder + Knee + Elbow + Ankle + Hip)
8.	A8	Active movement
II. Color		
9.	B1	Whole body blue or pale
10.	B2	Face pink; rest blue
11.	B3	Face, trunc pink; rest blue
12.	B4	Face, trunc, upper arms pink; rest blue
13.	B5	Face, trunc, upper arm, thigh pink; rest blue
14.	B6	Face, trunc, upper arm, thigh, four arm pink; rest blue
15.	B7	Whole body pink; hands and feet blue
16.	B8	Whole body pink; nails of fingers and toes blue

17.	B9	Whole body pink; nails of fingers blue
18.	B10	Whole body pink; nails of toes blue
19.	B11	Whole body pink
III. Respiratory Effort		
20	C1	None
21	C2	Rate abnormal character gasping
22	C3	Rate normal character gasping
23	C4	Rate abnormal character irregular
24	C5	Rate normal character irregular
25	C6	Rate abnormal character rhythmic
26	C7	Rate normal character rhythmic
IV. Heart Rate		
27	D1	None
28	D2	Below 100
29	D3	Normal (100 - 140)
V. Reflex Stimulation		
30	E1	No response
31	E2	Grimace but no sneeze
32	E3	First grimace & then sneeze
33	E4	First Sneeze & then grimace
34	E5	First touch sneeze

Observed set of features leading to the definite known cases of illness have been supplied to the system sequentially as first cum first serve basis to train the system. After each entry, an iterative evaluation process as described in section 2 starts to generate the value of RFW of each feature corresponding to each category. After each entry, the system updates the values of RFWs and tests for optimization with $\delta \sim 10^{-3}$. A positive result of optimization testing indicates that no more field cases are required to train the system. The system architecture is presented in figure 10.2.

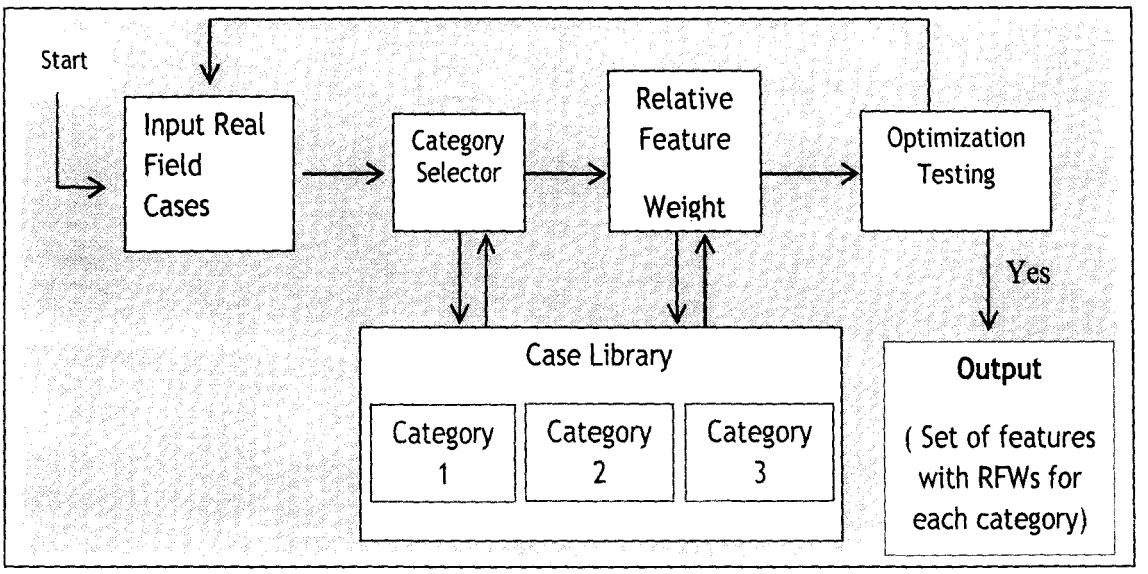

Fig. 10.2. System Architecture for feature selection.

10.4. Reasoning aspects

10.4.1. Category

The structure of the category (case table) in the case library has been designed using the category-exemplar model of case representation [46]. This essentially means that each case should have a set of features as depicted in fig. 10.3. The case memory is embedded in a linked structure of categories (case tables), cases and index pointers. Cases are associated with categories and are composed off various case descriptors (features). Neighbour cases differ from each other in one or small number of features. Different features are assigned different importance (Relative Feature Weights) in describing a case. An index may point to a case or case table.

10.4.2. Category selector

The category selector contains the indexed tables of the feature sets to the cases/categories. Depending upon the observed case(s) leading to definite goal(s) the category selector points the most appropriate case table(s) in the case library and incorporates the new case(s) in the corresponding case tables(s). The Relative Feature Weight Generator then becomes activated.

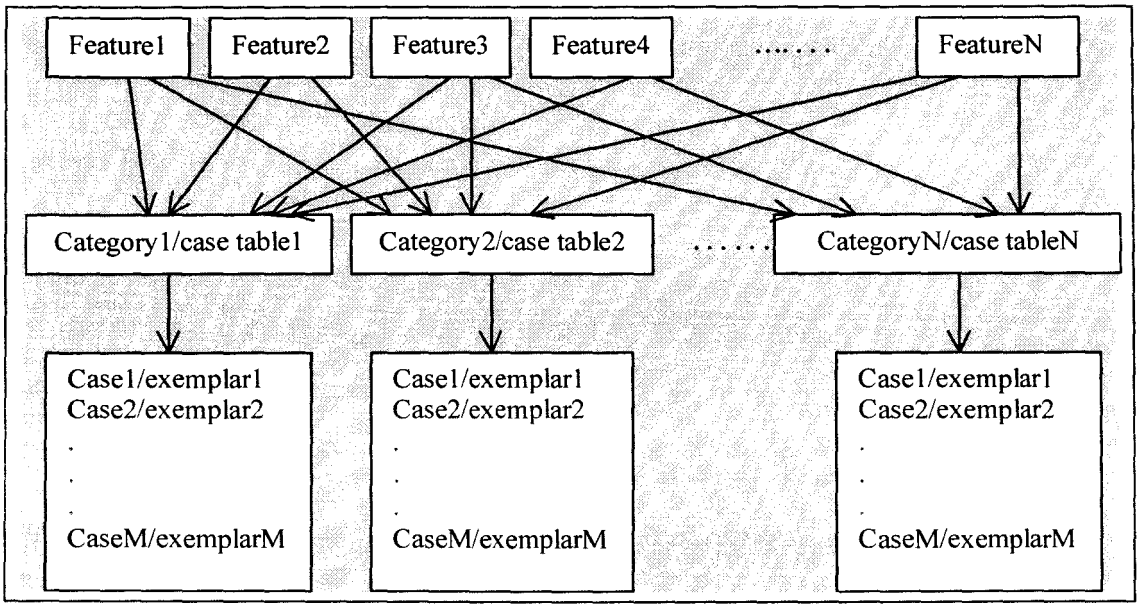

Fig. 10.3. Structure of features, categories and exemplars.

10.4.2.1. Relative feature weight generator

The Feature Weight Generator scans the corresponding case table(s) in which the new observed case(s) has been incorporated and generates the Relative Feature Weight (belief) of the features corresponding to the case(s), through an iterative and inductive process described in section 10.2. The generated Relative Feature Weight of each feature is then the subject of optimization testing.

10.4.3. Optimization testing

Optimization testing indicates the end of training. When sufficient numbers of definite observed cases are inserted in the case table, the Relative Feature Weight of each feature approaches towards stability; that means the system becomes optimized. We consider here a threshold value δ for measuring optimization. The relative feature weights of a category are tested for optimization with interval of 10 cases. We consider it optimized when

$$|\text{RFW m th} - \text{RFW (m+10) th}| \leq \delta$$

for all contributing features of the category. This interval may be accepted on the argument that the contributions from less than 10 cases might not be significant for optimization testing and on the worst case this may increase case memory by 10 cases only.

10.5. Theoretical foundation of the classifier

10.5.1. Feature vector (F)

Feature Vector (F) is a set of elements transformed from the feature set of all possible features in a domain. If N is the total number features in the feature set of a domain then the F can be represented as:

$$[F] = \begin{bmatrix} F_1 \\ F_2 \\ . \\ . \\ F_N \end{bmatrix} \tag{10.4}$$

where each of the elements F_1, F_2, ...F_N possess a value either 1 or 0 depending upon the set of observed features. If for a particular observation, the i-th feature is observed, then $F_i = 1$ otherwise it is 0. As an example, if a domain consists of total 6 features and for a particular case, features f_1, f_3 and f_6 are observed, then the feature set of these 6 features are transformed to a Feature Vector as:

$$[F] = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

10.5.2. Bias vector (Φ)

Bias is a critical barrier of the weighted sum above which a case is considered to be fired and the fixation of the bias is a vital task to improve the performance of the system. In our system, the bias for cases under a category has been fixed on the basis of Relative Feature Weights (RFW). We have classified the features of each category in 3 sub-sets; (i) weekly-relevant (wr), (ii) relevant (r) and (iii) strongly relevant (sr). The features having RFW greater than the arithmetic mean (M_a) are labeled as 'strongly-relevant' features. The features having RFW in the range between $M_a/2$ to M_a is the 'relevant' one and features those having RFW less than $M_a/2$ are the 'weekly-relevant'.

The rationale for estimation of bias value is based on the theme that at least one member of each group should have the contribution and the features with highest RFW within each sub-set should have the highest priority to be considered in fixing the bias. The bias for i-th category is

$$\Phi_i = [\mathrm{Max}(RFW_{wr}) + \mathrm{Max}(RFW_r) + \mathrm{Max}(RFW_{sr})]_i \qquad (10.5)$$

The Bias Vector is a set where the elements represent the biases (case firing threshold) of all the possible categories in the domain. For the present problem, the total number of categories in this domain is 3, so the Bias Vector (Φ) can be represented as:

$$[\Phi] = \begin{bmatrix} \Phi_1 \\ \Phi_2 \\ \Phi_3 \end{bmatrix} \qquad (10.6)$$

10.5.3. Synapse matrix (ω)

Synapse Matrix (ω) is a matrix whose elements are the Relative Feature Weights at saturation of all the features constructing the possible categories in a domain. After training of the system, the significant and non-significant features are identified and the corresponding Relative Feature Weights of all the categories are known. These Relative Feature Weights corresponding to a category constitutes the row of this Synapse Matrix. An element of the row will be considered as 0 (zero) if the corresponding feature has no contribution (non-significant feature) to that particular category. For the present problem, ω is a 3x34 matrix because the number of categories is 3 and the total number of features in a domain is 34. The ω can be represented as:

$$[\omega] = \begin{bmatrix} \omega_{1,1} & \omega_{1,2} & \omega_{1,3} & \cdots & \omega_{1,34} \\ \omega_{2,1} & \omega_{2,2} & \omega_{2,3} & \cdots & \omega_{2,34} \\ \omega_{3,1} & \omega_{3,2} & \omega_{3,3} & \cdots & \omega_{3,34} \end{bmatrix} \qquad (10.7)$$

Where ω_{11} is the Relative Feature Weight (synapse) of feature-1 of the category-1, ω_{12} is that of feature-2 of the category-1 and so on.

10.5.4. Activation value (X) and the system output (Y)

The general model of the classifier consists of a summing part followed by an output part. In this case, we have considered the classifier as a single layer neural network of McCulloch-Pitts (MP) model of neurons [48]. The number of neurons or nodes in the network is equal to the number of possible categories in the problem domain. The Activation Value is given by a weighted sum of input values and a bias term. The output signal is typically a nonlinear transfer function of the activation value. If $F_1, F_2, F_3. . . .F_N$ be the elements of the Feature Vector (F) and $\omega_{i1}, \omega_{i2}, \omega_{i3}, ...\omega_{iN}$ are the synapse values for i-the node, then the Activation Value (X_i) for i-th node can be represented as:

$$X_i = \sum_{j=1}^{N} \omega_{ij} . F_j - \Phi_i \qquad (10.8)$$

Where, N is the total number of features in the domain.

Then the output value of the i-th node can be defined by the output function

$$Y_i = f(x_i)$$

The transfer function, that we have used is a binary function, where

$$f(x_i) = 1, \quad \text{if } x > 0$$
$$= 0, \quad \text{if } x \leq 0$$

If $f(x_i) = 1$, the i-th category is considered to be fired; otherwise not considered. The McCulloch-Pitts model for i-th neuron is shown in figure 10.4.

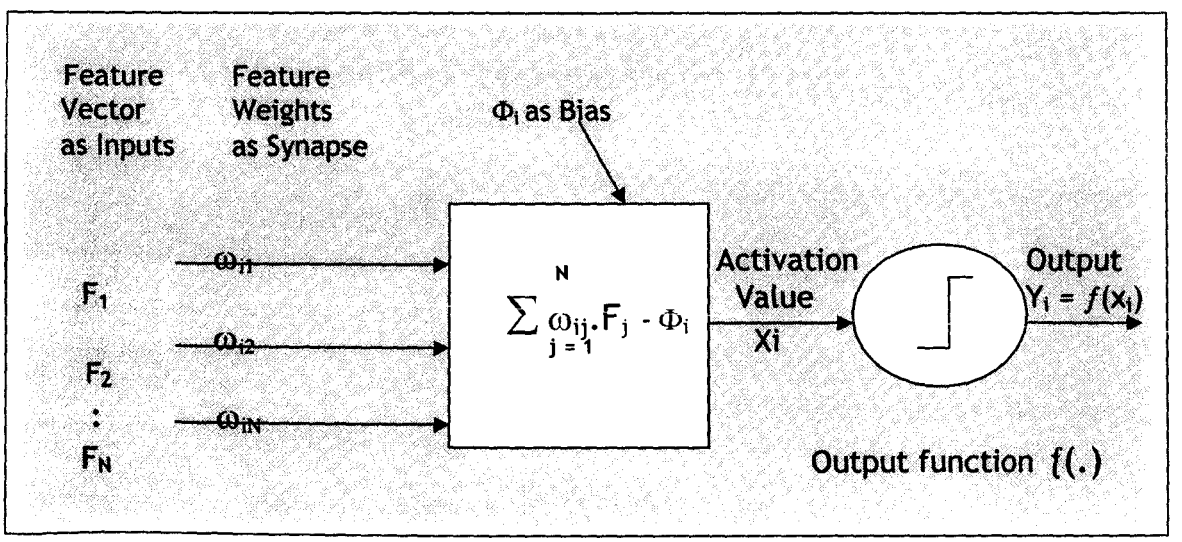

Fig. 10.4. McCulloch-Pitts model of a neuron.

10.6. Classifier system architecture

After the training, when a new case appears, the system receives the set of observed features of the new case and the Feature Vector Generator transforms observed feature set to Feature Vector (F). The system activates the Synapse Matrix Generator. The Synapse Matrix Generator then scans the case library to pick up the Relative Feature Weights of all the features for all categories and generates the Synapse Matrix (ω) according to equation (10.7). The Bias Vector Generator calculates the bias of each category and constructs the Bias Vector (Φ).

The classifier calculates the activation values (X) of the new case against each category and generates the outputs (Y) by using the transformation function. The categories for which the output is 1 are the classified category that is, the new case is of the type of this category. The block diagram of the system architecture is presented in figure 10.5.

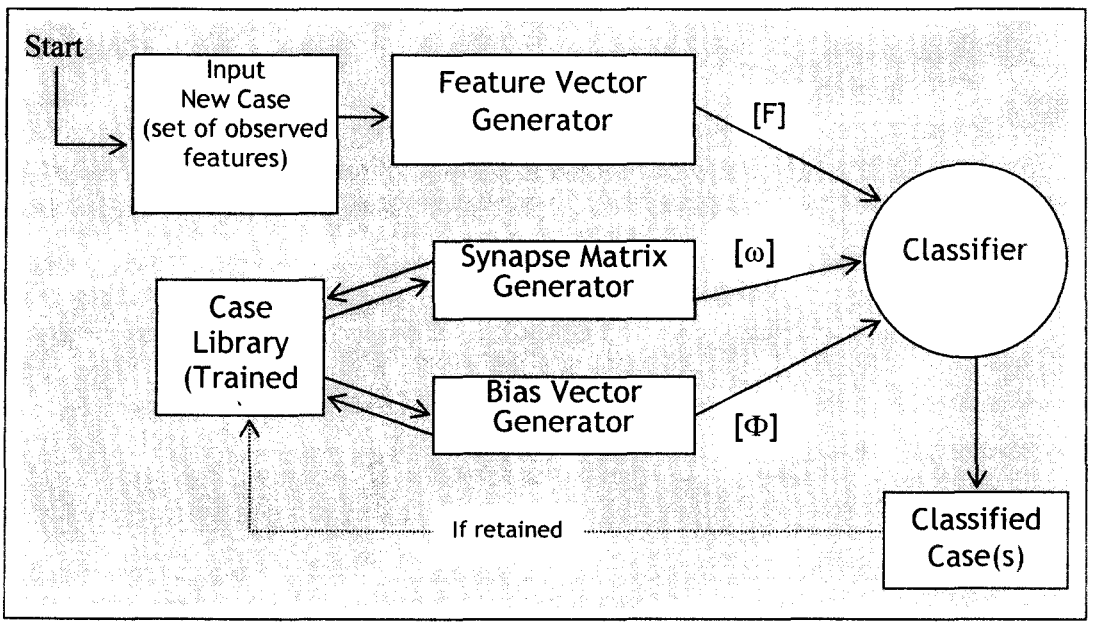

Fig. 10.5. Block diagram of the system architecture.

The classifier is a single layer neural network of McCulloch-Pitts model of neurons. For the present domain, the number of possible categories is 3, so the classifier contains 3 neurons (nodes). Each neuron is responsible to classify a particular category. The block diagram of the classifier is presented in figure 10.6.

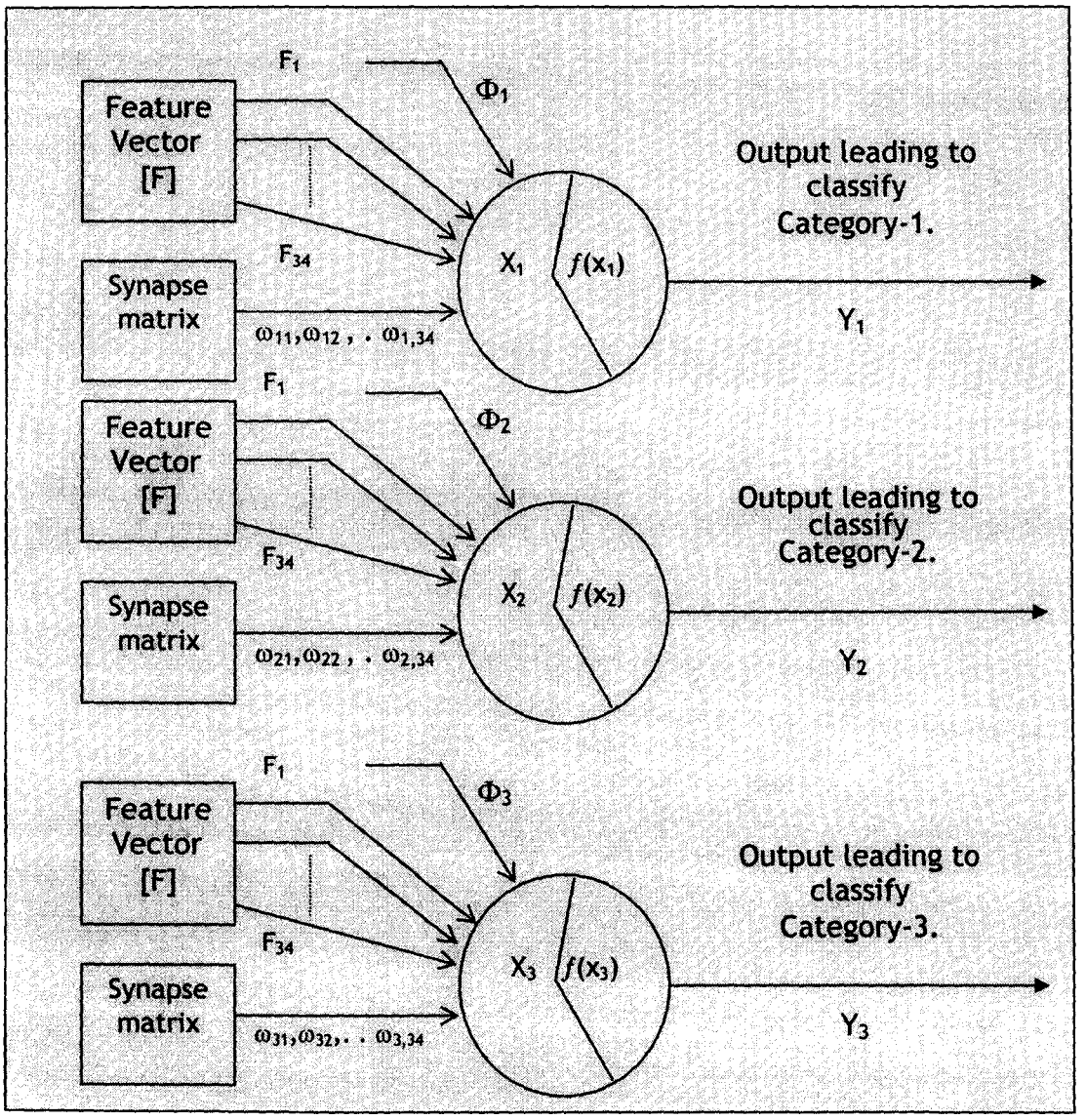

Fig. 10.6. Block diagram of the classifier.

Each neuron of this neural network accepts Feature Vector, Bias Vector and Synapse Matrix as inputs. The neuron then generates the Activation Value and gives output by using transfer function. Depending upon the output value, the case is considered to be classified.

Functionality of the proposed classifier architecture is illustrated in the following steps:

Step 1: Input set of observed features for new case.

Step 2: Generate feature vector from the observed set of features.

Step 3: Generate Synapse Matrix by using case library.

Step 4: Generate Bias Vector by using case library.

Step 5: Calculate the Activation Value of new case against each case.

Step 6: Calculate the Output against each category by using transfer function.

Step 7: If Output = 1 for any category, then the case is classified to that category.

10.7. Case illustrations

At the end of the feature selection process, as a result, the significant features corresponding to each category are identified by the system. The features obtained with RFW # 0 are the significant features and that of with RFW = 0 are the non-significant or noisy features. RFW = 0 indicates that, the noisy features have no contribution to a particular category and can be discarded.

The Relative Feature Weights of the significant features obtained against each category constitutes the non-zero elements of the Synapse Matrix. The identified significant features and RFWs (elements of Synapse Matrix) at optimization for each type of illness are presented in tables (10.3 - 10.5).

Table10.3. Identified significant features with their code and RFWs (elements of synapse matrix) for Type-1 illness.

Feature codes	Significant features	RFWs (non-zero elements of synapse matrix)
A1	Muscle Tone of Limbs: Flaccid	$\omega_{1,01} = 0.220$
A3	Muscle Tone of Limbs: Flexion at two joins and rest flaccid (Hip + Shoulder)	$\omega_{1,03} = 0.011$
B3	Color: Face, trunc pink; rest blue	$\omega_{1,11} = 0.094$
B5	Color: Face, trunc, upper arm, thigh pink; rest blue	$\omega_{1,13} = 0.070$
C1	Respiratory Effort: None	$\omega_{1,20} = 0.026$
C3	Respiratory Effort: Rate normal character gasping	$\omega_{1,22} = 0.011$
C4	Respiratory Effort: Rate abnormal character irregular	$\omega_{1,23} = 0.024$
C6	Respiratory Effort: Rate abnormal character rhythmic	$\omega_{1,25} = 0.022$
D2	Heart Rate: Below 100	$\omega_{1,28} = 0.325$
E1	Reflex Stimulation: No response	$\omega_{1,30} = 0.172$
E2	Reflex Stimulation:Grimace but no sneeze	$\omega_{1,31} = 0.024$

Table 10.4. Identified significant features with their code and RFWs (elements of synapse matrix) for Type-2 illness.

Feature codes	Significant features	RFWs (non-zero elements of synapse matrix)
A5	Muscle Tone of Limbs: Flexion at four joins and rest flaccid (Hip + Shoulder + Knee + Elbow)	$\omega_{2,05} = 0.132$
A8	Muscle Tone of Limbs: Active movement	$\omega_{2,08} = 0.050$
B2	Color: Face pink; rest blue	$\omega_{2,10} = 0.050$
B5	Color: Face, trunc, upper arm, thigh pink; rest blue	$\omega_{2,13} = 0.035$
B6	Color: Face, trunc, upper arm, thigh, four arm pink; rest blue	$\omega_{2,14} = 0.030$
C2	Respiratory Effort: Rate abnormal character gasping	$\omega_{2,21} = 0.130$
C7	Respiratory Effort: Rate abnormal character irregular	$\omega_{2,26} = 0.050$
D3	Heart Rate: Normal (100 - 140)	$\omega_{2,29} = 0.340$
E1	Reflex Stimulation: No response	$\omega_{2,30} = 0.142$
E3	Reflex Stimulation: First grimace & then sneeze	$\omega_{2,32} = 0.040$

Table 10.5. Identified significant features with their code and RFWs (elements of synapse matrix) for Type-3 illness.

Feature codes	Significant features	RFWs (non-zero elements of synapse matrix)
A5	Muscle Tone of Limbs: Flexion at four joins and rest flaccid (Hip + Shoulder + Knee + Elbow)	$\omega_{3,05} = 0.023$
A8	Muscle Tone of Limbs: Active movement	$\omega_{3,08} = 0.184$
B8	Color: Whole body pink; nails of fingers and toes blue	$\omega_{3,16} = 0.170$
B10	Color: Whole body pink; nails of toes blue	$\omega_{3,18} = 0.025$
C5	Respiratory Effort: Rate normal character irregular	$\omega_{3,24} = 0.027$
C6	Respiratory Effort: Rate abnormal character rhythmic	$\omega_{3,25} = 0.074$
C7	Respiratory Effort: Rate abnormal character irregular	$\omega_{3,26} = 0.018$
D3	Heart Rate: Normal (100 - 140)	$\omega_{3,29} = 0.330$
E1	Reflex Stimulation: No response	$\omega_{3,30} = 0.023$
E4	Reflex Stimulation: First Sneeze & then grimace	$\omega_{3,33} = 0.008$
E5	Reflex Stimulation: First touch sneeze	$\omega_{3,34} = 0.118$

10.7.1. Case study for Type-I Illness

The values of RFWs of significant features (presented in codes) obtained against increased number of retained cases (NRC) for type-I Illness category is presented in table 10.6. as a representative one.

Table10.6. The significant features and their RFWs for type-I Illness.

NRC	Relative Feature Weights of significant features of Type-1 Illness										
	A1	A3	B3	B5	C1	C3	C4	C6	D2	E1	E2
1	0.198	0.000	0.198	0.000	0.000	0.000	0.198	0.000	0.198	0.000	0.198
2	0.283	0.000	0.071	0.071	0.071	0.000	0.071	0.000	0.283	0.071	0.071
3	0.293	0.000	0.128	0.032	0.032	0.000	0.032	0.032	0.293	0.128	0.032
4	0.188	0.023	0.082	0.082	0.023	0.023	0.023	0.023	0.335	0.188	0.023
5	0.202	0.014	0.052	0.115	0.052	0.014	0.014	0.014	0.317	0.202	0.014
6	0.216	0.009	0.077	0.077	0.034	0.009	0.009	0.034	0.310	0.216	0.009
7	0.237	0.008	0.106	0.058	0.025	0.008	0.025	0.025	0.320	0.164	0.025
8	0.237	0.004	0.122	0.043	0.019	0.004	0.019	0.043	0.310	0.175	0.019
9	0.197	0.016	0.101	0.064	0.016	0.016	0.016	0.036	0.326	0.197	0.016
10	0.210	0.013	0.118	0.052	0.013	0.013	0.029	0.029	0.329	0.161	0.029
11	0.215	0.011	0.130	0.042	0.011	0.011	0.024	0.042	0.321	0.170	0.024
12	0.223	0.009	0.109	0.056	0.021	0.009	0.021	0.035	0.320	0.180	0.021
13	0.230	0.008	0.123	0.048	0.018	0.008	0.030	0.030	0.321	0.154	0.030
14	0.236	0.006	0.133	0.040	0.015	0.006	0.040	0.027	0.322	0.133	0.040
15	0.213	0.013	0.119	0.053	0.013	0.013	0.036	0.023	0.332	0.147	0.036
16	0.216	0.011	0.128	0.046	0.011	0.011	0.031	0.031	0.327	0.154	0.031
17	0.221	0.011	0.113	0.055	0.018	0.011	0.028	0.028	0.326	0.162	0.028
18	0.226	0.009	0.122	0.049	0.016	0.009	0.036	0.025	0.326	0.144	0.036
19	0.230	0.008	0.109	0.057	0.023	0.008	0.032	0.023	0.324	0.152	0.032
20	0.234	0.007	0.098	0.066	0.029	0.007	0.029	0.020	0.323	0.158	0.029
21	0.238	0.007	0.106	0.059	0.026	0.007	0.036	0.019	0.324	0.144	0.036
22	0.221	0.011	0.098	0.068	0.025	0.011	0.033	0.017	0.330	0.153	0.033
23	0.223	0.010	0.105	0.062	0.022	0.010	0.031	0.022	0.327	0.158	0.031
24	0.226	0.009	0.095	0.069	0.028	0.009	0.028	0.020	0.325	0.163	0.028
25	0.230	0.009	0.103	0.063	0.026	0.009	0.034	0.019	0.326	0.151	0.034
26	0.215	0.012	0.096	0.071	0.024	0.012	0.031	0.017	0.331	0.159	0.031
27	0.218	0.011	0.088	0.076	0.029	0.011	0.029	0.016	0.329	0.163	0.029
28	0.221	0.010	0.094	0.071	0.027	0.010	0.027	0.021	0.327	0.167	0.027
29	0.208	0.014	0.088	0.077	0.025	0.014	0.025	0.019	0.330	0.173	0.025
30	0.210	0.013	0.082	0.082	0.030	0.013	0.023	0.018	0.328	0.176	0.023
31	0.213	0.012	0.087	0.077	0.028	0.012	0.022	0.022	0.327	0.180	0.022
32	0.216	0.011	0.092	0.072	0.026	0.011	0.026	0.020	0.328	0.170	0.026
33	0.219	0.011	0.097	0.067	0.024	0.011	0.024	0.024	0.326	0.172	0.024
34	0.208	0.014	0.092	0.073	0.023	0.014	0.023	0.023	0.329	0.178	0.023
35	0.211	0.013	0.098	0.069	0.022	0.013	0.027	0.022	0.330	0.169	0.027
36	0.213	0.013	0.101	0.065	0.021	0.013	0.026	0.026	0.329	0.171	0.026

37	0.215	0.012	0.096	0.069	0.024	0.012	0.024	0.024	0.327	0.174	0.024
38	0.218	0.011	0.100	0.065	0.023	0.011	0.028	0.023	0.328	0.166	0.028
39	0.219	0.011	0.104	0.062	0.022	0.011	0.026	0.026	0.326	0.168	0.026
40	0.210	0.013	0.100	0.066	0.020	0.013	0.025	0.025	0.329	0.173	0.025
41	0.212	0.012	0.103	0.063	0.019	0.012	0.024	0.028	0.327	0.175	0.024
42	0.214	0.012	0.098	0.067	0.022	0.012	0.022	0.026	0.326	0.178	0.022
43	0.217	0.011	0.102	0.064	0.021	0.011	0.026	0.026	0.327	0.170	0.026
44	0.219	0.011	0.097	0.067	0.024	0.011	0.024	0.024	0.326	0.172	0.024
45	0.220	0.010	0.093	0.071	0.027	0.010	0.023	0.023	0.325	0.175	0.023
46	0.222	0.010	0.096	0.068	0.026	0.010	0.026	0.022	0.326	0.168	0.026
47	0.215	0.012	0.093	0.072	0.025	0.012	0.025	0.022	0.328	0.172	0.025
48	0.216	0.012	0.096	0.069	0.024	0.012	0.024	0.024	0.327	0.174	0.024
49	0.217	0.011	0.092	0.072	0.027	0.011	0.023	0.023	0.326	0.176	0.023
50	0.220	0.010	0.095	0.069	0.026	0.010	0.026	0.022	0.327	0.169	0.026
51	0.221	0.010	0.091	0.072	0.028	0.010	0.025	0.021	0.326	0.171	0.025
52	0.222	0.010	0.088	0.075	0.031	0.010	0.024	0.020	0.324	0.173	0.024
53	0.223	0.009	0.090	0.072	0.030	0.009	0.022	0.022	0.324	0.175	0.022
54	0.216	0.011	0.088	0.076	0.029	0.011	0.022	0.022	0.326	0.179	0.022
55	0.217	0.011	0.084	0.078	0.031	0.011	0.021	0.021	0.325	0.180	0.021
56	0.218	0.010	0.087	0.075	0.030	0.010	0.020	0.023	0.324	0.182	0.020
57	0.221	0.010	0.090	0.073	0.029	0.010	0.023	0.023	0.324	0.176	0.023
58	0.221	0.010	0.092	0.070	0.028	0.010	0.022	0.025	0.324	0.178	0.022
59	0.215	0.011	0.090	0.073	0.027	0.011	0.021	0.024	0.325	0.181	0.021
60	0.218	0.011	0.093	0.071	0.026	0.011	0.023	0.023	0.326	0.175	0.023
61	0.218	0.011	0.095	0.069	0.025	0.011	0.022	0.025	0.325	0.177	0.022
62	0.219	0.010	0.092	0.071	0.027	0.010	0.022	0.024	0.324	0.179	0.022
63	0.221	0.010	0.095	0.069	0.027	0.010	0.024	0.024	0.325	0.173	0.024
64	0.223	0.010	0.097	0.067	0.026	0.010	0.026	0.023	0.325	0.168	0.026
65	0.218	0.011	0.095	0.070	0.025	0.011	0.025	0.022	0.327	0.171	0.025
66	0.218	0.011	0.097	0.067	0.024	0.011	0.024	0.024	0.326	0.173	0.024
67	0.219	0.010	0.094	0.070	0.026	0.010	0.024	0.024	0.325	0.174	0.024
68	0.221	0.010	0.096	0.068	0.026	0.010	0.026	0.023	0.326	0.169	0.026
69	0.222	0.010	0.094	0.070	0.027	0.010	0.025	0.022	0.325	0.171	0.025
70	0.223	0.009	0.091	0.072	0.029	0.009	0.024	0.021	0.325	0.172	0.024
71	0.224	0.009	0.093	0.070	0.028	0.009	0.026	0.021	0.325	0.168	0.026
72	0.219	0.011	0.091	0.073	0.028	0.011	0.025	0.020	0.327	0.170	0.025
73	0.220	0.010	0.093	0.071	0.027	0.010	0.025	0.022	0.326	0.172	0.025
74	0.221	0.010	0.090	0.073	0.029	0.010	0.024	0.022	0.325	0.173	0.024
75	0.223	0.010	0.093	0.071	0.028	0.010	0.026	0.021	0.326	0.169	0.026

The relative feature weights plotted against the number of cases retained (RFW-curve) are shown in the graphs (figs. 10.7 - 10.17).

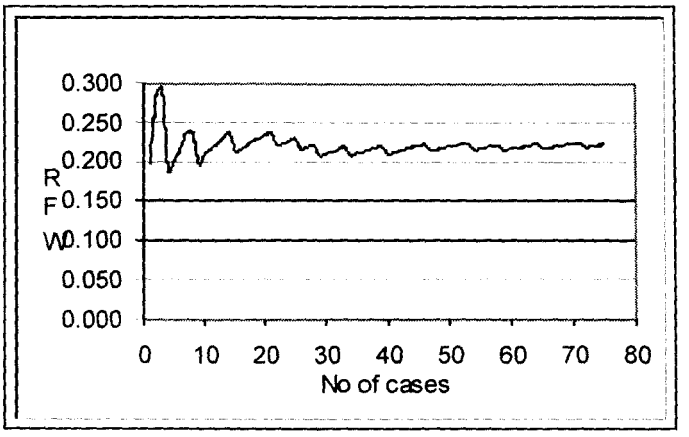

Fig.10.7. RFW curve of Feature A1 of Type-I illness.

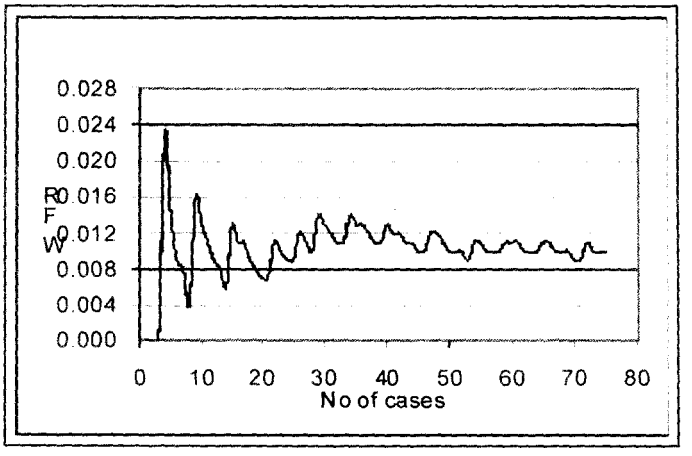

Fig.10.8. RFW curve of Feature A3 of Type-I illness.

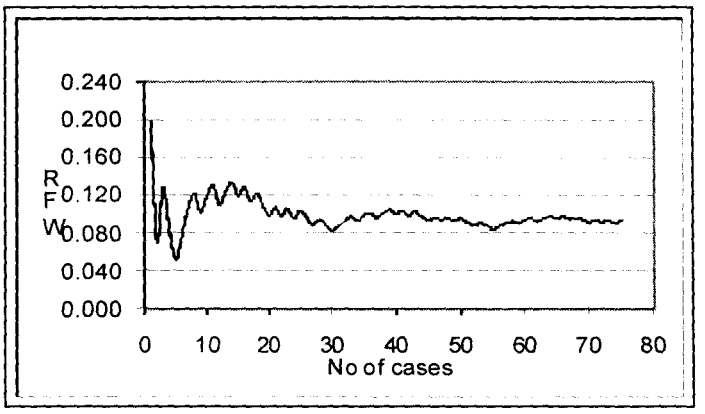

Fig.10.9. RFW curve of Feature B3 of Type-I illness.

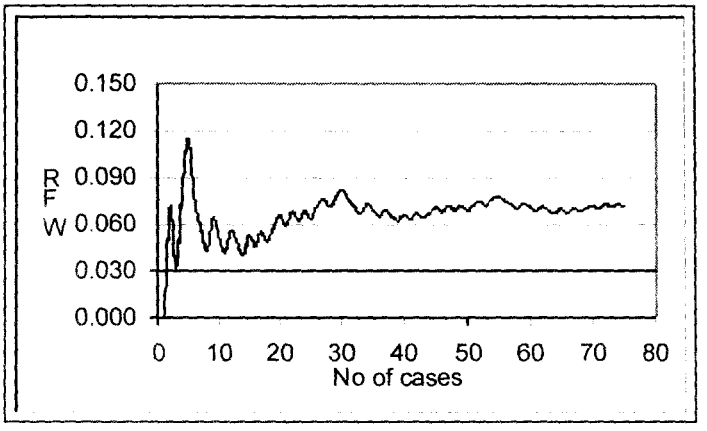

Fig.10.10. RFW curve of Feature B5 of Type-I illness.

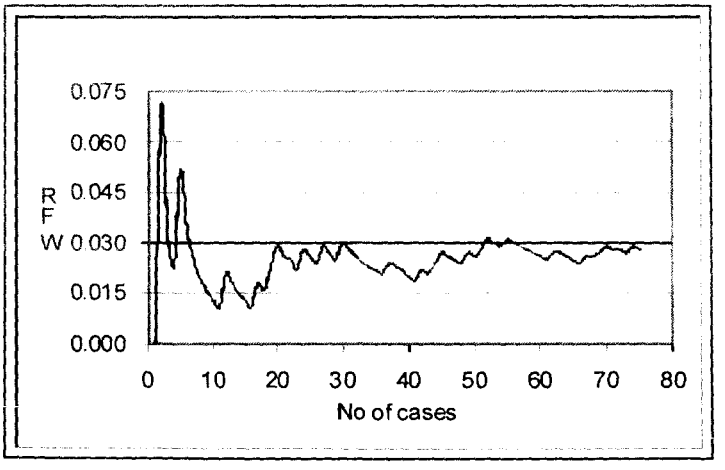

Fig.10.11. RFW curve of Feature C1 of Type-I illness.

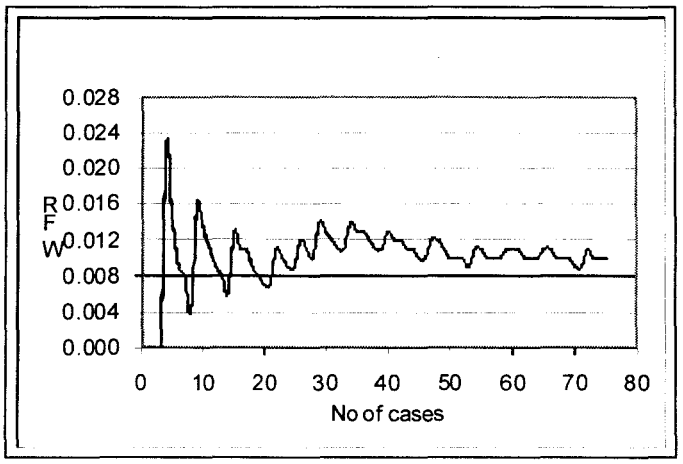

Fig.10.12. RFW curve of Feature C3 of Type-I illness.

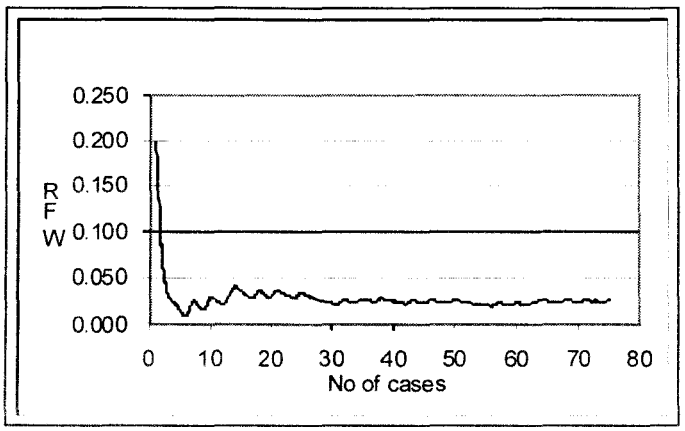

Fig.10.13. RFW curve of Feature C4 of Type-I illness.

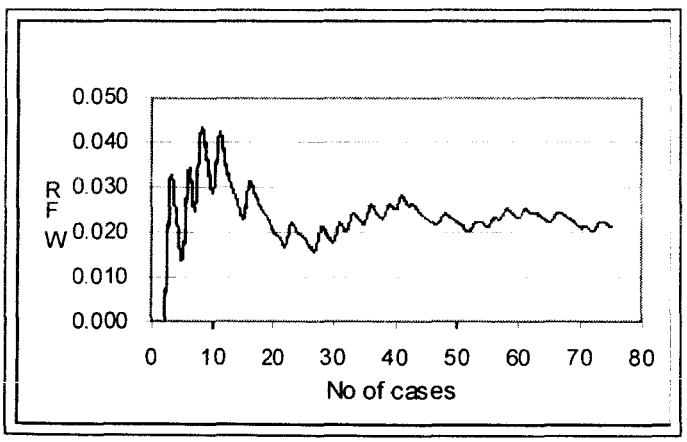

Fig.10.14. RFW curve of Feature C6 of Type-I illness.

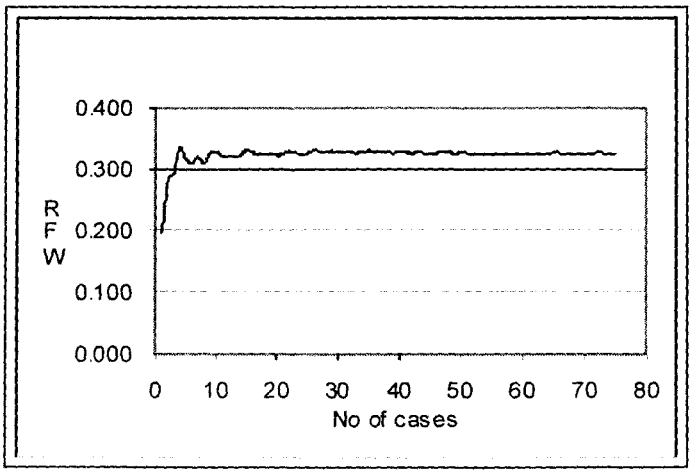

Fig.10.15. RFW curve of Feature D2 of Type-I illness.

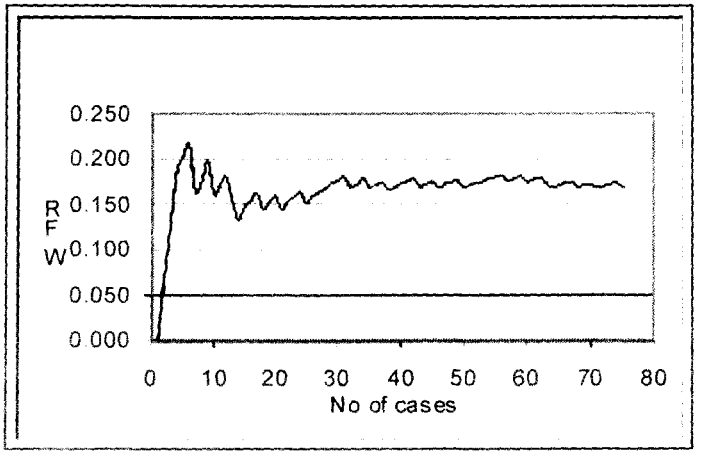

Fig.10.16. RFW curve of Feature E1 of Type-I illness.

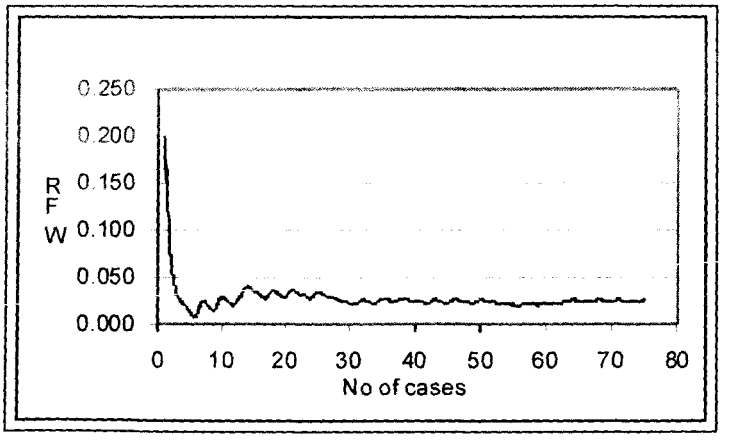

Fig.10.17. RFW curve of Feature E2 of Type-I illness.

Now let us proceed to calculate the bias values (elements of Bias Vector) for each category. As a primary task to calculate the bias, the type of the features (i.e. Strongly relevant, relevant or weekly relevant) are identified based on arithmetic mean and the elements of bias vector have been calculated by using equation (10.5) and presented in the table 10.7.

Table 10.7. Elements of Bias Vector.

Category	Type of Illness	Bias values
Category-1	Type-1	$\Phi_1 = 0.421$
Category-2	Type-2	$\Phi_2 = 0.430$
Category-3	Type-3	$\Phi_3 = 0.431$

Now the system is ready to classify the new cases.

To test the performance of the system, more than 40 definite and real field cases of the problem domain were supplied to the system to classify the cases. The system classified all of them accurately. Two examples are presented here.

Case example 1.

The observed set of features and feature codes for case example 1 are:

1. Muscle Tone of Limbs: Active movement (A8)
2. Color: Whole body pink; nails of toes blue (B10)
3. Respiratory Effort: Rate abnormal character rhythmic (C6)
4. Heart Rate: Normal (100 - 140) (D3)
5. Reflex Stimulation: First touch sneeze (E5)

Case example 2.

The observed set of features and feature codes for case example 2 are:

1. Muscle Tone of Limbs: Flaccid (A1)
2. Color: Face, trunc, upper arm, thigh pink; rest blue (B5)
3. Respiratory Effort: Rate abnormal character rhythmic (C6)
4. Heart Rate: Below 100 (D2)
5. Reflex Stimulation: No response (E1)

On the basis of the observed feature sets, the classifier generates the activation values of the above cases in respect of all the categories and uses the transfer

function to classify the cases of specific categories. The activation values and the classification of cases by the system is shown in tables 10.8 and 10.9.

Table 10.8. System's output verses actual field observation for Case example 1.

Categories	$\Sigma\,\omega.\,F$	Φ	X	Y	Classification by the system	Actual field observation
1. Type-1	0.022	0.421	-0.399	0		Type - 3
2. Type-2	0.390	0.430	-0.040	0		
3. Type-3	0.731	0.431	+0.300	1	Type - 3	

Table 10.9. System's output verses actual field observation for Case example 2.

Categories	$\Sigma\,\omega.\,F$	Φ	X	Y	Classification by the system	Actual field observation
1. Type-1	0.809	0.421	+0.388	1	Type - 1	Type - 1
2. Type-2	0.177	0.430	-0.253	0		
3. Type-3	0.097	0.431	-0.334	0		

10.8. Conclusion and discussion

Our system sorts out the significant features along with their relative feature weights and the non-significant features are ruled out after having the field cases during training. The learning is robust in the sense that it avoids irrelevant storage. The method is computationally efficient in polynomial order in its learning when provided with finite number of training examples. This also helps in estimating the number of retained cases for a problem domain. This estimate should be useful for the development of a case-based knowledge system for the domain of interest. Here it has been assumed that the features in the feature set are mutually independent.

From the above examples, it is evident that the proposed approach might be considered as a potential approach for classifier type of problems using CBR. As the system gets trained with real field cases, no heuristics or approximate estimation is concerned and thereby reducing the uncertainty to a great extent. The only uncertainty, if any, is involved with the real field data for the new cases. It is a hybrid system of case-based reasoning and classifier approach. The classifier removes the problems associated with retrieval and adaptation of CBR and CBR eliminates the problems concerned with the training of a classifier. It is a system where the merits of CBR and classifier have been clubbed.

References

1. W. E. Nelson. Text Book of Pediatrics; 15th edition, Book 1, W. B. Saunders Company; Bangalore, 1996.
2. O. D. Saugstad. Practical aspects of resuscitating asphyxiated newborn infants. Eur J Pediatr. 157 (suppl 1). S11-S15. 1998.
3. C. Palme-Kilander. Methods of resuscitation in low-Apgar-score newborn infants: a national survey. Acta Paediatr. 81, pp.739 -744, 1992.
4. World Health Report. Geneva, Switzerland: World Health Organization; 1995.
5. Adapted from International Guidelines for Neonatal Resuscitation: An Excerpt From the Guidelines 2000 for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care: International Consensus on Science, Contributors and Reviewers for the Neonatal Resuscitation Guidelines Pediatrics,2000;106;e29. http://www.pediatrics.org/cgi/content/full/106/3/e29. Accessed on 16/06/2006.
6. J. L. Kolodner and R. M. Kolodoner. Using experience in clinical problem solving: introduction and framework; IEEE Trans Syst. Man Cybernet, 17, pp. 420-431, 1987.
7. E. Bareiss, B. Porter and C. Wier. PROTOS: An exempler-based learning apprentice; Int. J. Man-Machine Studies, 29, pp. 549-561, 1988.
8. J. Berger. ROENTGEN: A case-based approach to radiation therapy planning. Proc. of a Workshop on Case-Based Reasoning. Morgan Kaufmann; San Mateo, CA. pp. 218-223. 1989.
9. Y. Jang. HYDI: A Hybrid System with Feedback for Diagnosing Multiple Disorders; PhD. Thesis (TR-576), Computer Science, Massachusetts Institute of Technology, MA. 1993.
10. C. Bradbum. FLORENCE: Synthesis of case-based and model-based reasoning in nursing care planning system, Comp. Nursing, 11, pp. 20-24, 1993.
11. R. T. Macura and K. Macura. Case-based reasoning: opportunities and applications in health care, Artificial Intelligence in Medicine, 9, pp. 1-4, 1997.
12. Li D. Xu. Developing a case-based knowledge system for AIDS prevention, Expert Systems, 11(4), pp. 237-244, 1994.

13. M. Sebban and R. Nock. A hybrid filter/wrapper approach of feature selection using information theory, Pattern Recognition, 35, pp. 835-846, 2002.

14. A. Y. Ng. On feature selection - learning with exponentially many irrelevant features as training examples, Proc. of the Fifteenth Int. Conf. on Machine Learning, Madison, WI, pp. 404- 412, 1998.

15. R. O. Duda, P. E. Hart, and D. G. Stork. Pattern Classification, John Wiley and Sons, Second edition, 2001.

16. M. Dash and H. Liu. Feature selection for classification, Intelligent Data Analysis, vol 1, pp. 131-156, 1997.

17. I. Guyon and A. Elisseeff. An Introduction to Variable and Feature Selection, Journal of Machine Learning Research, 3, pp. 1157-1182, 2003.

18. J. G. Dy and C. E. Brodley. Feature subset selection and order identification for unsupervised learning, Proc. of the Seventeenth Int. Conf. on Machine Learning, Stanford, CA., pp. 247-254, 2000.

19. H. Almuallin and T. G. Dietterich. Learning with many irrelevant features, In Ninth National Conf. on Artificial Intelligence, pp. 547-552, MIT Press, 1991.

20. K. Kira and L. A. Rendell. The feature selection problem : traditional methods and a new algorithm, Proc. of the Tenth Nat. Conf. on Artificial Intelligence, San Jose, CA, AAAI Press / The MIT Press, Menlo Park, CA., pp. 129-134, 1992.

21. A. L. Blum and P. Langley. Selection of relevant features and examples in machine learning, Artificial Intelligence, pp. 245-271, 1997.

22. S. Das. Filters, wrappers and a boosting-based hybrid for feature selection, Proc. of the 18th Int. Conf. on Machine Learning, Williamstown, MA, pp. 74-81, 2001.

23. M. Dash and H. Liu. Feature selection for clustering, Proc. of the Fourth Pacific Asia Conf. on Knowledge Discovery and Data Mining, (PAKKD-2000), Kyoto, Japan, Springer, Berlin, pp. 110-121, 2000.

24. R. Kohavi and G. H. John. Wrappers for feature subset selection. Artificial Intelligence 97 (12), pp. 273-324. 1997.

25. H. Liu, H. Motoda, and L. Yu. Feature selection with sampling, Proc. of the Nineteenth Int. Conf. on Machine Learning, Sidney, Australia, pp. 395-402, 2002.

26. H. Liu, H. Motoda, and L. Yu. A selective sampling approach to active feature selection, Artificial Intelligence, 159, pp. 49-74, 2004.

27. L. Li and L. X. Knowledge based problem solving: an approach to health assessment, Expert Systems with Applications, 16, pp. 33-42, 1999.

28. L. Talavera. Feature selection as a preprocessing step for hierarchical clustering, Proc. of Int. Conf. on Machine Learning (ICML99), Bled, Slovenia, pp. 389-397, 1999.

29. K. D. Ashley and E. L. Rissland. Waiting on Weighting: A symbolic least commitment approach, in Proceedings of the Seventh National Conf. on Artificial Intelligence, St. Paul, MN: Morgan Kaufmann, pp. 239-244, 1988.

30. D. W. Aha and R. L. Goldstone. Concept learning and Flexible Weighting, Proc. of the Fourteenth Annual Conf. of the Cognitive Science Society. Bloomington, IN, Laurence Erlbaum, pp. 534-539, 1992.

31. D. Wettschereck and D. W. Aha. Weighting feature, Proc. of the 1st Int. Conf. on Case-Based Reasoning, pp. 347-358, 1995.

32. D. Wettschereck, T. Mohori and D. W. Aha. A review and empirical evaluation of feature weighting methods for a class of lazy learning algorithms, Artificial Intelligence Review, 1997.

33. I. Ghosh and R. K. Samanta. A Model for Case-Based Learning, The Journal of Computer Society of India, vol. 30, no. 4, pp. 9-14, 2000.

34. N. K. Bose and P. Liang. Neural Network Fundamentals with Graphs, Algorithms and Applications. McGraw-Hill, Int. Editions. 1996.

35. M. H. Hassoun. Fundamentals of Artificial Neural Networks. Cambridge, MA: MIT Press. 1995.

36. M. Sarkar and B. Yegnanarayana. An evolutionary programming-based probabilistic neural network construction technique, in Proc. of IEEE Int. Conference on Neural Networks. (Houston, USA), June 9-12, 1997.

37. R. M. Welch, S. K. Sengupta, A. K. Goroch, P. Rabindra, N. Rangaraj, M. S. Navar. Polar cloud and surface classification using AVHRR imagery: An inter comparison of methods, Journal of Applied Meteorology, 31, pp. 405-420, 1992.

38. P. Domingos and M. Pazzani. Beyond independence: Conditions for the optimality of the simple Bayesian classifier, Proceedings of the 13th International Conference on Machine Learning, San Francisco, CA. Morgan Kaufmann, pp. 105-112, 1996.

39. B. V. Dasarathy. Nearest Neighbour (NN) Norms: NN Pattern Classification Techniques, Los Alamitos, California: IEEE Computer Society Press, 1991.

40. P. Langley and W. Iba. Average-case Analysis of a Nearest Neighbour Algorithm, Proceedings of the 13th International Joint Conference on Artificial Intelligence (IJCAI), San Mateo, CA: Morgan Kaufmann, pp. 889-894, 1993.

41. D. Michie, D. J. Spiegelhalter and C. C. Taylor. Machine Learning, Neural and Statistical Classification, Eds., UK: Ellis Horwood Ltd., 1994.

42. J. Rachlin, S. Kasif, S. Salzberg, and D. W. Aha. Towards a Better Understanding of Memory-Based Reasoning Systems, Proceedings of the 11th International Machine Learning Conference (ML94), San Francisco, CA: Morgan Kaufmann, 1994.

43. D. J. C. MacKay. Bayesian methods for supervised neural networks. The Handbook of Brain Theory and Neural Networks, (M. A. Arbib, ed.), Cambridge, MA: MIT Press, pp.144-149, 1995.

44. D. Lowe. Radial basis function networks, The Handbook of Brain Theory and Neural Networks, (M. A. Arbib, ed.), Cambridge, MA: MIT Press, pp. 779-782, 1995.

45. T. R. Payne and P. Edwards. Implicit Feature Selection with the Value Difference Metric, 13th European Conference on Artificial Intelligence (ECAI-1998), Ed. Henri Prade, John Wiley & Sons. Ltd, 1998.

46. Agnar Aamodt and Enric Plaza. Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches, AICom - Artificial Intelligence Communications, IOS Press, Vol. 7, No. 1, pp. 39-59, 1994.

47. M. G. Goswami. An Approach towards the development of an expert system for pediatric problem domain, Ph. D. Thesis, University of North Bengal, India, 1999.

48. B. Yegnanarayana. Artificial Neural Networks. Prentice-Hall of India Pvt. Ltd., New Delhi, pp. 26-27, 1999.