PREDICTION OF NEONATAL JAUNDICE USING FUZZY CLUSTERING METHODS

Hadi Talebian Moghaddam*, Farzad Towhidkhah
Department of Bioelectrical Engineering
Biomedical Engineering Faculty
Amirkabir University of Technology, No. 424, Hafez Ave., PO Box 15875-4413, Tehran, Iran
htmoghaddam@yahoo.com, towhidkhah@aut.ac.ir

Rasoul Khayati
Department of Biomedical Engineering
Engineering Faculty
Shahed University, Tehran, Iran
khayati@shahed.ac.ir

Mohammad Torkaman, Zohreh Kavehmanesh
Department of Pediatrics
Bagiyatallah Medical Sciences University, Tehran, Iran
mhmmdtrkmmn@yahoo.com, z_kaveh@hotmail.com

ABSTRACT
Prediction of diseases in most cases is still challenging and unresolved by physicians. Jaundice is the most common and one of the most vexing problems that can occur in the newborns. Although most jaundiced infants are otherwise perfectly healthy, they make us anxious because bilirubin is potentially toxic to the central nervous system and kernicterus can occur.

Some different fuzzy clustering methods including fuzzy c-means (FCM), subtractive clustering (SC) and fuzzy adaptive resonance theory mapping (fuzzy ARTMAP) are presented for prediction of the risk of jaundice before and after delivery (in first 48 hours) of newborns. A total of 552 medical records were collected from newborns during April to June 2006 in two general hospitals in Tehran, Iran. To evaluate results of the applied methods we used evaluation performance matrix criteria, which include correct classification (CC%), sensitivity (SE%), and specificity (SP%). The above mentioned criteria for jaundice prediction before delivery were approximately 76%, 97%, 56%, while for jaundice prediction after delivery were 81%, 88%, 67%, respectively. These results show that the proposed systems can achieve satisfying results for predicting risk of jaundice considering this fact that physicians do not have any estimation about probability of jaundice appearance.

KEY WORDS
Fuzzy Clustering, Neonatal Jaundice, Prediction, fuzzy ARTMAP, Fuzzy C-Means (FCM), Subtractive Clustering

1. Introduction
Jaundice is a common and, in most cases, benign problem in neonates. In 1990s Jaundice was the prevalent cause of returning of newborns to hospitals [1]. So, early prediction of newborns liable to Jaundice is important. Hemolytic disease of the newborn is a common cause of neonatal jaundice. Nonetheless, because of the immaturity of the pathways of bilirubin metabolism, many newborn infants without evidence of hemolysis become jaundiced. Nonetheless, untreated, severe indirect hyperbilirubinemia is potentially neurotoxic. Jaundice is observed during the 1st wk of life in approximately 60% of term infants and 80% of preterm infants. The unconjugated form of bilirubin is neurotoxic in infants at certain concentrations and under various conditions. Conjugated bilirubin is not neurotoxic but indicates a potentially, serious disorder.

Compared with adults, newborn infants have a twofold to threefold greater rate of bilirubin production (610 mg/kg/24 hr versus 3 mg/kg/24 hr). This is caused in part by an increased red blood cell mass (higher hematocrit) and a shortened erythrocyte life span of 70-90 days, compared with the 120day erythrocyte life span in adults. Risk factors for indirect hyperbilirubinemia include maternal diabetes, race, prematurity, drugs, male sex, oxytocin induction, breast-feeding, and a sibling who had jaundice. A family history of neonatal jaundice, exclusive breast-feeding, Asian race, and maternal age older than 25 year identify approximately 60% of cases of extreme hyperbilirubinemia [2,3]. In infants without these variables, indirect bilirubin levels rarely rise above 12 mg/dL, whereas infants with several risk factors are more likely to have higher bilirubin levels. Indirect bilirubin levels in full-term infants decline to adult levels (1 mg/dL) by 10-14 days of life. Prediction of which neonatal infants are at risk for exaggerated jaundice can be based on hour-specific bilirubin levels in the 1st 24-72 hr of life [3]. In general, a search to determine the cause of jaundice should be made if (1) it appears in the 1st 24-36 hr of life, (2) serum bilirubin is rising at a rate faster than 5
mg/dL/24 hr, (3) serum bilirubin is greater than 12 mg/dl, in full-term (especially in the absence of risk factors) or 10-14 mg/dL in preterm infants, (4) jaundice persists after 10-14 days of life, or (5) direct-reacting bilirubin is greater than 2 mg/dL at any time. Among other factors suggesting cause of jaundice are family histories of hemolytic disease, pallor, hepatomegaly, splenomegaly, failure of phototherapy to lower bilirubin [3].

In what follows, in section 2 comes a review of previous works on jaundice. Then a summary of the applied clustering methods is given in section 3. Section 4 includes data gathering and feature selection matters. Results and their Evaluation method come in section 5. Discussion and conclusion come at last in sections 6.

2. Literature Study

A model for jaundice treatment in newborns is the mathematical model presented by Charles Lee and his students [4]. Lee et al. proposed a model of bilirubin circulation in the body by use of mathematical equations.

A. Seyfang et al presented an algorithm for diagnosis and treatment of Jaundice in newborns by use of ASBRU language in 2002 [5].

American Academy of Pediatrics also suggested an algorithm for diagnosing and treatment of Jaundice [6].

A research in Switzerland in 2000 shows the effect of season concerning the amounts of bilirubin in 5540 newborns cured by phototherapy during 1993-1996 Which says, the maximum number of phototherapies occurred during the months May to August [7].

David Stevenson et al suggested the use of End-tidal carbon monoxide (CO) for ambient CO (ETCOc), i.e., test of expiration carbon monoxide to force jaundice in newborns, considering release of carbon monoxide when heme changes to bilirubin. See Fig. 1 [8].

Khayat's research said about different ways of jaundice treatment defining different states and relations between them using Stateflow and Simulink toolboxes of MATLAB software [2].

In 2003 Ahouraei et al tried to diagnose neonatal jaundice using a neural network. The neural network used in their study was a standard feed forward network with two hidden layer, trained with back propagation [10].

3. Clustering Methods

The objective of cluster analysis is the classification of objects according to similarities among them, and organizing of data into groups. Clustering techniques are among the unsupervised methods; they do not use prior class identifiers. The main potential of clustering is to detect the underlying structure in data, not only for classification and pattern recognition, but also for model reduction and optimization.

3.1. Fuzzy ARTMAP System

Detailed descriptions of the fuzzy ARTMAP system can be found in [11]. For completeness, in the following, we introduce only the necessary details. Each fuzzy ARTMAP system shown in Fig. 2 includes a pair of fuzzy ART modules, ARTa and ARTb, that are linked together via an inter-ART module, F_{ab}, called a map field. During supervised learning, ARTa receives an input in the complement code form \( \vec{I}^c = A^c = (\vec{A}^c, \vec{d}^c) \), and ARTb receives an input also in the complement code, \( \vec{I}^T = B^T = (\vec{b}^T, \vec{l}^T) \). Note that each component in \( I \) is in the interval \([0,1]\] and complement coding is a normalization rule that preserves amplitude information. The map field is used to form predictive associations between categories and to realize the match-tracking rule.

It ensures autonomous system operation in real time and works by increasing the vigilance parameter \( \rho \) of ARTa. Parameter \( \rho \) calibrates the minimum confidence that ARTa must have in a recognition category, or hypothesis, activated by an input A in order for ARTa to accept that category, rather than search for a better one through an automatically controlled process of hypothesis testing.
The lower the value of $\rho_a$ is the larger the number of categories is. A predictive failure of at ART$_b$ increases $\rho_a$ by the minimum amount needed to at ART$_a$, using a mechanism called match tracking. Hypothesis testing leads to the selection of a new ART$_a$ category, which focuses attention on a new cluster of A that is better able to predict B. Owing to these mechanisms, fuzzy ARTMAP systems becomes one of a rapidly growing family of incremental learning pattern recognition systems.

### 3.2. The Fuzzy C-means Algorithm

The fuzzy C-means method is the most widely used algorithm to detect constrained fuzzy c-partitions. A constrained fuzzy c-partition can be briefly described as follows: Let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of n unlabeled feature data vectors with $x_k \in \mathbb{R}^p$ ($1 \leq k \leq n$), and $c$ ($2 \leq c \leq n$) be a number of fuzzy subsets (clusters) defined in $X$. Given that the membership function of the $k$th vector to the $i$th cluster is denoted as $u_{ik} = \{u_i(x_k), 1 \leq i \leq c, 1 \leq k \leq n\}$, the c fuzzy clusters constitute a constrained fuzzy c-partition in $X$ if the next three conditions are satisfied,

$$0 \leq u_{ik} \leq 1, \ \forall i, k \quad (1a)$$

$$0 \leq \sum_{k=1}^{n} u_{ik} \leq n, \ \forall i \quad (1b)$$

$$\sum_{i=1}^{c} u_{ik} = 1, \ \forall k \quad (1c)$$

Whenever the last condition, (1c), is not satisfied the fuzzy c-partition is said to be unconstrained. Fuzzy c-means is able to detect constrained fuzzy c-partitions by minimizing an objective function [12]. This optimization procedure is described as follows:

$$\text{Minimize} \left\{ J_m(U,V;X) = \sum_{k=1}^{n} \left[ \sum_{i=1}^{c} (u_{ik})^m \|x_k - v_i\|^2 \right] \right\}$$

subject to equation 1(c)

where $U = \{[u_{ik}], 1 \leq i \leq c, 1 \leq k \leq n\}$ is the partition matrix, $V = \{[v_i], 1 \leq i \leq c\}$ with $v_i \in \mathbb{R}^p$ is the vector of the resulted cluster centers, $m \in (1, \infty)$ is a factor to adjust the membership degree weighting effect, and $\| \cdot \|_k$ is any inner product norm. The cluster centers and the respective membership functions that solve the constrained optimization problem in (2) are given by the following equations:

$$v_i = \frac{\sum_{k=1}^{n} u_{ik}^m x_k}{\sum_{k=1}^{n} u_{ik}^m}, \quad 1 \leq i \leq c \quad (3)$$

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left( \frac{\|x_k - v_i\|^2}{\|x_k - v_j\|^2} \right)^{2/m-1}}, \quad 1 \leq i \leq c, 1 \leq k \leq n \quad (4)$$

Eqs. (3) and (4) constitute an iterative optimization procedure. There are two types of iteration algorithms: the first type, for the $k$th iteration, calculates in sequence the quantities $V_{k-1}, U_{k-1}, V_k$, and validates the condition $\|V_k - V_{k-1}\|_o \leq \epsilon$. The second type calculates the sequence $U_{k-1}, V_k, U_k$, and checks the condition $\|U_k - U_{k-1}\|_o \leq \epsilon$. The main difference between the two algorithms is that the second one is slower, since more parameters have to converge [13]. Recently, Cheng et al. improved the speed of this type by a factor of 2–3 times [14].

### 3.3. Subtractive Clustering (SC)

In this paper, Chiu’s subtractive clustering is applied which possesses some interesting advantages, especially in a neuro-fuzzy identification context [15]. In fact, subtractive clustering is an efficient algorithm, which does not require optimization, being for this reason a good choice for the initialization of neuro-fuzzy networks. Fuzzy c-means and other optimization-based clustering techniques would lead to excessive computer work because they perform an unnecessary optimization phase prior to network training. Also, progressive clustering and compatible cluster merging algorithms are computationally expensive and need metrics for validation of individual clusters. Therefore, despite their potential, they are too complex for a simple initialization of a fuzzy neural network [3]. Chiu’s algorithm belongs to the class of potential function methods, being, a variation of the mountain method. In this class of algorithms, a set of points are defined as possible group centers, each of them being interpreted as an energy source. In subtractive clustering the center candidates are the data samples themselves, which overcomes the main limitation of the mountain method. In fact, there, the candidates are defined in a grid, leading to “curse of dimensionality” problems.

Let $Z^N$ be a set of N data samples, $z_1, z_2, \ldots, z_N$, defined in an $m + n$ space, where $m$ denotes the number of inputs and $n$ the number of outputs. In order to make the range of values in each dimension identical, the data samples are normalized, so that they are limited by a hypercube. As it was referred, it is admitted that each of the samples defines a possible cluster center. Therefore, the potential associated to $z_i$ is (5):

$$P_i(z_i, Z^N) = \sum_{j=1}^{N} e^{-a||z_i - z_j||^2}, \quad \alpha = \frac{4}{r_a^2}, i = 1, 2, \ldots, N \quad (5)$$

where $r_a > 0$ is the radii parameter, a constant which defines the neighborhood radius of each point. Thus, points $z_i$ located out of the radius of $z_i$ will have a smaller influence in its potential. On the other hand, the effect of points close to $z_i$ will grow with the proximity. In this way, points with a dense neighborhood will have higher associated potentials. After computing the potential for each point, the one with the highest potential is selected as the first cluster center. Next, the potential of all the
remaining points is reduced. Defining \( z^*_i \) as the first group center and denoting its potential as \( P^*_i \), the potential of the remaining points is reduced as in (6):

\[
P_i \leftarrow P_i - P_i^* e^{-\beta|z_i - z_i^*|^2}, \quad \beta = \frac{4}{r_b^2}
\]

(6)

where the constant \( r_b > 0 \) defines the neighborhood radius with sensitive reductions in its potential. In this way, points close to the selected center will have their potentials reduced in a more significant manner, and so the probability of being selected as centers diminishes. This procedure has the advantage of avoiding the concentration of identical clusters in denser zones. Therefore, \( r_b \) is selected in order to be slightly higher than \( r_s \), to avoid closely spaced clusters. Typically, \( r_b = 1.5 r_s \).

After performing the reduction of potential for all of the candidates, the one with the highest potential is selected as the second cluster. Then, the potential of the remaining points is again reduced. Generically, after determining the \( r \)th group, the potential is reduced as follows (7):

\[
P_i \leftarrow P_i - P_i^* e^{-\beta|z_i - z_i^*|^2}
\]

(7)

The procedure of center selection and potential reduction is repeated until stopping criterion mentioned in Table 1 is reached.

**Table 1. Stopping Criterion for Subtractive Clustering**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_i^* &gt; \varepsilon^{up} P_i )</td>
<td>Accept ( z_i^* ) as the next cluster center and continue</td>
</tr>
<tr>
<td>Otherwise, ( P_i^* &gt; \varepsilon^{down} P_i^* )</td>
<td>Reject ( z_i^* ) and finish the algorithm.</td>
</tr>
<tr>
<td>Otherwise</td>
<td>Let ( d_{\text{min}} ) be the shortest distance between ( z_i ) and all the centers already found</td>
</tr>
<tr>
<td>( d_{\text{min}} / r_s + P_i^* / P_i^* \geq 1 )</td>
<td>Accept ( z_i^* ) as the next cluster center and continue</td>
</tr>
<tr>
<td>Otherwise</td>
<td>Reject ( z_i^* ) and assign it the potential 0.0.</td>
</tr>
</tbody>
</table>

Select the point with higher potential as the new \( z_i^* \).

Repeat the test.

End If

End If

End If

In Table 1, \( \varepsilon \) specifies a threshold above which the point is selected as a center with no doubts and \( \varepsilon \) specifies the threshold below which the point is definitely rejected. The third case is where the point is characterized by a good tradeoff between having a sufficiently high potential and being distant enough from the clusters determined before. Typically, \( \varepsilon^{up} = 0.5 \) and \( \varepsilon^{down} = 0.15 \).

As it can be understood from the description of the algorithm, the number of clusters to obtain is not pre-specified. However, it is important to note that the radii parameter is directly related to the number of clusters found. Thus, a small radius will lead to a high number of rules, which, if excessive, may result in overfitting. On the other hand, a higher radius will lead to a smaller number of clusters, which may originate underfitting, and so, models with reduced representation accuracy.

In practice it is necessary to test several values for radii and select the most adequate according to the results obtained. However, despite the fact that some radii values should be tested, this parameter gives an initial hint on the number of clusters necessary. This constitutes an important advantage over optimization based and other classes of clustering algorithms, when little information is known regarding the best number of clusters. Another advantage of subtractive clustering is that the algorithm is noise robust, since outliers do not significantly influence the choice of centers, due to their low potentials [3].

## 4. Materials and Methods

### 4.1. Data Acquisition

Clinical data about newborns were collected by filling approved blank sheets for taking information & data from delivery occurrences during three months from two general hospitals in Tehran. Method of data acquisition was based on clinical data, physicians' prescriptions and questionnaire from the newborn’s family. This data includes both newborns affected by jaundice and the healthy ones. At the end of data gathering period we had 552 files subject to the research. Because of some missing values for each input variable we had to remove some of incomplete record files. Of course because of different aspects for prediction of jaundice, numbers of removals were different for each model. Finally 515 and 333 files were used for jaundice prediction before and after delivery, respectively. Some statistical information about gathered data are presented in Table 2.

**Table 2. Statistical Information about Gathered Database**

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Total research cases (Healthy + Jaundiced)</th>
<th>Total Jaundice cases of the research</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>552</td>
<td>391</td>
</tr>
<tr>
<td>Prediction Data</td>
<td>515</td>
<td>333</td>
</tr>
</tbody>
</table>

### 4.2. Feature Selection and Applied Models

There is a wide range of factors that affect neonatal bilirubin levels [16]. Some of these factors have been identified only in large epidemiological studies and their clinical relevance is questionable, but there are some that have been shown repeatedly to have an important influence on total serum bilirubin levels as a measure of jaundice severity [17]. These factors can be mentioned as
follows: genetic, ethnic and familial influences, maternal factors, events during labor and delivery, neonatal factors. Table 3 refers to epidemiology of neonatal jaundice [9]. We had to ignore some of mentioned factors in our current study, because of some restrictions in recording.

Table 3. Epidemiology of neonatal jaundice [9]

<table>
<thead>
<tr>
<th>Associated Factors</th>
<th>Effect on Neonatal Serum Bilirubin Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>Increase, Decrease</td>
</tr>
<tr>
<td>Genetic or Familial Maternal</td>
<td>Previous sibling with jaundice* Diabetes* Hypertension</td>
</tr>
<tr>
<td>Drugs Administered to Mother</td>
<td>Oxytocin* Phenobarbital</td>
</tr>
<tr>
<td>Infants</td>
<td>Low birth weight, Decreasing gestation Male gender* Breast-feeding*</td>
</tr>
<tr>
<td>Other</td>
<td>Short hospital stay after birth*</td>
</tr>
</tbody>
</table>

*Most common clinically important factors

Hence, selected features for our systems are as follows.

Features used for prediction before delivery include mother’s O blood group, mother’s having diabetes, mother’s having hypertension, fetus' gender (males have more risk of jaundice), being the first child of family, siblings' jaundice background and mother’s age. It is assumed that fetus's gender has been declared by ultrasonography. Features used for early prediction after delivery include ABO incompatibility, Rh incompatibility, Cesarean delivery, mother's regional (spinal) anesthesia for delivery, oxytocin induction to mother, mother's hypertension, mother's diabetes, infant's gender, being the first child, prematurity, siblings' jaundice background, mother's age, infant's weight, gestational age. The clustering methods for prediction before and after delivery designed using fuzzy clustering toolbox of MATLAB software. Three clustering methods (SC, FCM and fuzzy ARTMAP) used for this purpose. To choose the best architecture of each clustering method, it was trained and tested for different parameters, e.g. cluster radius for SC method, number of clusters for FCM method and vigilance value for fuzzy ARTMAP method. The best results for each model and relevant configuration will come in results section.

5. Results and Evaluation Method

One of the most common evaluators is called evaluation matrix. Four combinations of classifier output and desired output are possible for jaundice prediction, which is shown in Table 4. Based on these parameters some criteria are calculated which follow. In these formulae TP, FP, FN, and TN are number of occurrence of corresponding state [18].

Table 4. Classifier Performance Evaluation Matrix

<table>
<thead>
<tr>
<th>Jaundice Presence</th>
<th>YES</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positive</td>
<td>TP</td>
<td>FN</td>
</tr>
<tr>
<td>False Positive</td>
<td>FP</td>
<td>TN</td>
</tr>
</tbody>
</table>

Table 5. Results of prediction before delivery

<table>
<thead>
<tr>
<th>Clustering Method</th>
<th>Time (sec)</th>
<th>Train CC%</th>
<th>Train SE%</th>
<th>Train SP%</th>
<th>Test CC%</th>
<th>Test SE%</th>
<th>Test SP%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC</td>
<td>0.12</td>
<td>75.38</td>
<td>97.12</td>
<td>53.56</td>
<td>76.03</td>
<td>96.52</td>
<td>55.73</td>
</tr>
<tr>
<td>FCM</td>
<td>4.42</td>
<td>64.05</td>
<td>66.14</td>
<td>59.07</td>
<td>70.97</td>
<td>73.64</td>
<td>64.36</td>
</tr>
<tr>
<td>ARTMAP</td>
<td>65.04</td>
<td>80.07</td>
<td>78.69</td>
<td>83.37</td>
<td>64.44</td>
<td>80.98</td>
<td>53.45</td>
</tr>
</tbody>
</table>

Table 6. Results of prediction after delivery

<table>
<thead>
<tr>
<th>Clustering Method</th>
<th>Time (sec)</th>
<th>Train CC%</th>
<th>Train SE%</th>
<th>Train SP%</th>
<th>Test CC%</th>
<th>Test SE%</th>
<th>Test SP%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC</td>
<td>0.98</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>81.27</td>
<td>88.36</td>
<td>66.57</td>
</tr>
<tr>
<td>FCM</td>
<td>6.18</td>
<td>98.59</td>
<td>99.64</td>
<td>95.73</td>
<td>78.42</td>
<td>85.10</td>
<td>65.948</td>
</tr>
<tr>
<td>ARTMAP</td>
<td>0.9</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>75.03</td>
<td>80.63</td>
<td>67.84</td>
</tr>
</tbody>
</table>

In a suitable experiment the sensitivity, specificity and correct classification can be defined as follows.

Sensitivity (SE): TP/(TP+FN). It means the probability that a test result is positive given the subject has the disease.

Specificity (SP): TN/(TN+FP). It means the probability that a test result is negative given a subject does not have the disease.

Correct Classification (CC): (TP+TN)/(TP+FP+TN+FN). It means the probability that the test result reflects the true disease state. Probability of a correct classifier output is estimated as the proportion of cases for which the classifier output is correct. The performance of a classifier should, wherever possible, be expressed in terms of sensitivity, specificity and correct classification.

Table 5 shows the results of prediction before delivery. Input data are in the form of 7 dimensional vectors, which were introduced in section 4.2. From the total 552 records 515 ones were without missing data, of which 362 data entries were used as training data and 153 data entries were used to test the network.

Table 6 shows the results of prediction after delivery. Input data are in the form of 15 dimensional vectors, which were introduced in feature selection section. From the total 552 records 333 ones were without missing data, of which 204 data entries were used as training data and 129 data entries were used to test the network.

6. Conclusion and Discussion

6.1. Conclusion

This idea that can we predict risk of neonatal jaundice
before and after delivery tried to be evaluated. For this purpose three clustering methods (i.e. SC, FCM and fuzzy ARTMAP) are presented for prediction of the risk of jaundice before and just after delivery of newborn. This research was done after collection of a total of 552 medical records from infants born during 12 April up to 15 June 2006 in two general hospitals in Tehran. Because of some missing values for each input variable, impossibility of complete follow up and communication problems we had to remove some of incomplete record files. Because of different aspects for prediction of jaundice, number of removals was different for each model. Finally 515 and 333 files were used for jaundice prediction before and after delivery, respectively.

The number of parameters in these models is different due to different available input parameters for each model. To evaluate results of these models we used evaluation performance matrix criteria, which include percentage of correct classification (CC%), sensitivity (SE%), and specificity (SP%). The above mentioned criteria for jaundice prediction before delivery were approximately 76%, 97%, 56%, while for jaundice prediction after delivery were 81%, 88%, 67%, respectively. These results show that the proposed clustering methods can achieve satisfying results for predicting risk of jaundice considering this fact that physicians do not have a clear estimation about probability of jaundice appearance.

6.2. Discussion

Use of jaundice prediction models will help to improve the health in society. This depends on modeling the complicated problems concerning jaundice in real situations. Although some directions and algorithms presented by society of physicians do not show these complications but they appear at clinical diagnosis. Modeling of these cases succeeds only when we have enough data in hand along with views of expert physicians. The different number of data entries used in each model is due to different number of missing data for each feature. Since number of features for prediction before and after delivery was 7 and 15, respectively; the total missing data for after delivery is more than before delivery. As can be seen results of prediction before delivery are not very satisfactory. This is due to insufficient information for this prediction. However considering this fact that there is not any clear method about jaundice prediction up to now by physicians, even this level of prediction which is gained by fuzzy clustering is at least a good step toward solving jaundice prediction problem. In case of prediction after delivery these results are more acceptable because some additional associated risk factors about the infant are available.

Acknowledgement

Authors acknowledge their best for guides and cooperation to Dr Talebian Moghaddam, Dr Soltanpour, Dr Ghadiani, Mrs Vatanparast, Naysabourou, Mr Ghielkhani, Management and Staff of Baqiyatallah & Najmiah hospitals and everyone who helped or encouraged the authors.

References

[5] Andreas Sayfang, Silvia Miksch, Mar Marcos, Combining Diagnosis and Treatment using Asbru.