ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.10

A Study on CTG Monitoring System Using Dataminig Techniques

Dr.J.K.Kanimozhi[1], R.Hemalatha[2]
Assitant Professor, PG[1], M.Phil Scholar, PG[2]
Department of Computer Science
Sengunthar Arts and Science College, Thiruchencode, India

Abstract:
Cardiotocography is one of the most widely used technique for recording changes in fetal heart rate (FHR) and uterine contractions. Assessing cardiotocography is crucial in that it leads to identifying fetuses which suffer from lack of oxygen, i.e. hypoxia. It is one of the most common diagnostic techniques to evaluate maternal and fetal well-being during pregnancy and before delivery. By observing the Cardiotocography trace patterns doctors can understand the state of the fetus. Even few decades after the introduction of cardiotocography into clinical practice, the predictive capacity of the existing methods remains inaccurate. In a previous work (Sundar.C and et al, 2016), we showed that a model based CTG data classification system using a supervised artificial neural network (ANN) can classify the CTG data better than most of the other methods. But, the performance of the normal neural network based classifier was limited because of the presence of potential outliers in the training data. The presence of outliers in training data affects the neural network training as well as testing. In this work, we present improved classification models which will consider outliers in the data and eliminate them from training phase of the classification process. The improved classifier was capable of identifying Normal, Suspicious and Pathologic condition with very good accuracy than normal methods.

Keywords: ANN, CTG, Fetal Distress, Fetal heart rate, IBM-NN, Outlier Detection

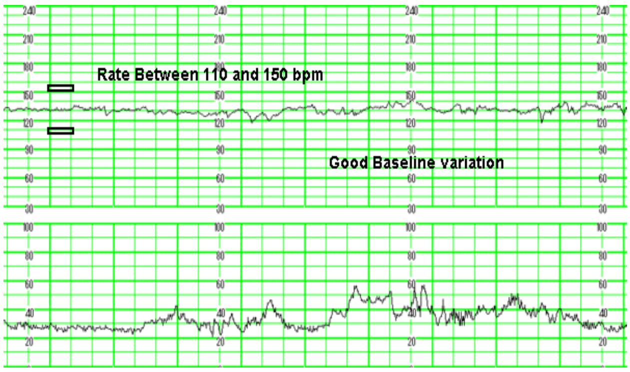
1. Introduction

One of the major challenges in medical domain is the extraction of comprehensible knowledge from medical diagnosis data such as Cardiotocography (CTG) data. In this information era, the use of machine learning tools in medical diagnosis is increasing gradually. This is mainly because the effectiveness in classification and recognition systems have improved to a great extend to help medical experts in diagnosing diseases.

Cardiotocography:
Cardiotocography (**CTG**) is a technical means of recording the fetal heartbeat and the uterine contractions during pregnancy. The machine used to perform the monitoring is called a **cardiotocograph,** more commonly known as an **electronic fetal monitor** (**EFM**). Fetal monitoring was invented by Doctors Bradfield, Orvan Hess and Edward Hon. CTG monitoring is widely used to assess fetal wellbeing. A review found that in the antenatal period (before labour) there is no evidence to suggest that monitoring women with high-risk pregnancies benefits the mother or baby although research around this is old and should be interpreted with caution.

The same review found that computerised CTG machines resulted in lower numbers of baby deaths than the traditional CTG machines. Cardiotocography or CTG is a test usually done in the third trimester of pregnancy. It is done to see if your baby's heart beats at a normal rate and variability. A CTG done in your third trimester is also known as a 'non stress test' because your baby is not under the 'stress' of labour. The correct utilization of this test can help prevent the baby from dying due to a shortage of oxygen. A cardiotocography is a technique that is used to monitor the heartbeat of the foetus along with keeping a check on uterine contractions. This can be deciphered from the name itself – cardio (heart) to co (uterine contractions) and graph (recording).

Internal Method:
Internal measurement of FHR involves attaching The device (an electrode) is placed close to the scalp of the foetus to get a reading. This form of monitoring is uncommon and is performed if there is difficulty in capturing the heartbeat in external monitoring. However, it is more accurate when compared to external monitoring and so may be the preferred method in complicated cases.



Baseline variability:
The amount in Beats per Minute (BPM) by which the baseline varies

Type 1	< 5 BPM	Abnormal
Type 2	5 – 10 BPM	Normal Sleep
Type 3	10 – 25 BPM	Normal Active
Type 4	> 25 BPM	Abnormal

External Method:
This is the most common form of testing and is done by placing the equipment on the abdomen of the mother. It consists of an elastic belt with two round plate the size of a cricket ball. One plate emits an ultrasound frequency which is used to detect the baby's heartbeat. The other plate is used to record the pressure on the mother's abdomen along with the

contractions. The belt is connected to a device that is used to read the signals that are being registered. Sometimes a gel may be applied to get a stronger signal. The reading by the device is given in the form of a loud beating sound. As some pregnant women can find this discomforting or distracting, there is a volume knob that can be used to lower the sound produced. Prior to going into labour, the mother would be asked to press a button on the machine each time she feels the baby move.

2. ANN Based Classification:

Here in this classification, use supervised learning by using a set of training data which is accompanied by class labels. When a new data arrives then classification of that data will be done based on the training set by generating descriptions of the classes. In addition to training set, also have a test data set that is used to determine the effectiveness of a classification. In general, commonly used and popular neural networks can be trained to recognize the data directly, whereas in simple networks there is a chance of the system being complex and training may be difficult. The time taken and the accuracy of classification depend on the dimension of the input given and also on the dimension in the training data. For input data with high dimension, the process will take a longer time.

Structuring the Network:

The number of layers and the number of processing elements per layer are important decisions. These parameters to a feed forward, back-propagation topology are also the most ethereal - they are the "art" of the network designer. There is no quantifiable, best answer to the layout of the network for any particular application.

There are only general rules picked up over time and followed by most researchers and engineers applying this architecture to their problems.

Rule One*:*

As the complexity in the relationship between the input data and the desired output increases, the number of the processing elements in the hidden layer should also increase.

Rule Two:

If the process being modeled is separable into multiple stages, then additional hidden layer(s) may be required. If the process is not separable into stages, then additional layers may simply enable memorization of the training set, and not a true general solution effective with other data.

Rule Three:

The amount of training data available sets an upper bound for the number of processing elements in the hidden layer(s). To calculate this upper bound, use the number of cases in the training data set and divide that number by the sum of the number of nodes in the input and output layers in the network. Then divide that result again by a scaling factor between five and ten. Larger scaling factors are used for relatively less noisy data. If you use too many artificial neurons the training set will be memorized. If that happens, generalization of the data will not occur; making the network useless on new data sets. A single-layer network of S logsig neurons having R inputs is shown below in full detail on the left and with a layer diagram on the right.

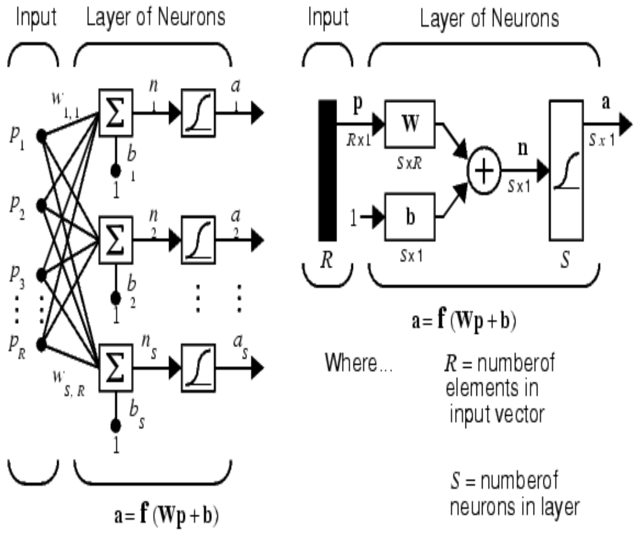


Figure 1: Feed forward Network

Feed forward networks often have one or more hidden layers of sigmoid neurons followed by an output layer of linear neurons. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors. The linear output layer lets the network produce values outside the range -1 to +1. On the other hand, if you want to constrain the outputs of a network (such as between 0 and 1), then the output layer should use a sigmoid transfer function.

3. The Improved BI_LEVEL Network Model:

In this work, the records in the training data classified with class labels. The Eigen vectors of the training data is used to reduce the dimension of the training data as well as testing data. Used the reduced dimensionality of training data as well as testing data were trained and classified. Outliers or abnormal records in the training data are detecting during the first stage of training and testing of the Back Propagation Neural network (BPN). After detecting outliers, those outliers will be removed from the training data, and again the same network will be trained with the outlier removed data to improve the training performance of the neural network and all the outliers will be included in the classification process. So, in this work, are going to address some of the machine learning based hybrid datamining techniques for the better classification of CTG data.

Advantages of IBM-NN Classification Model

Since the reduced dimension data is only used for training and testing, there will be considerable improvement in performance in terms of accuracy of classification.

Further the low dimensional data will also reduce the overhead in training. So that will get little improvement in performance than BM-NN model during practical implementations.

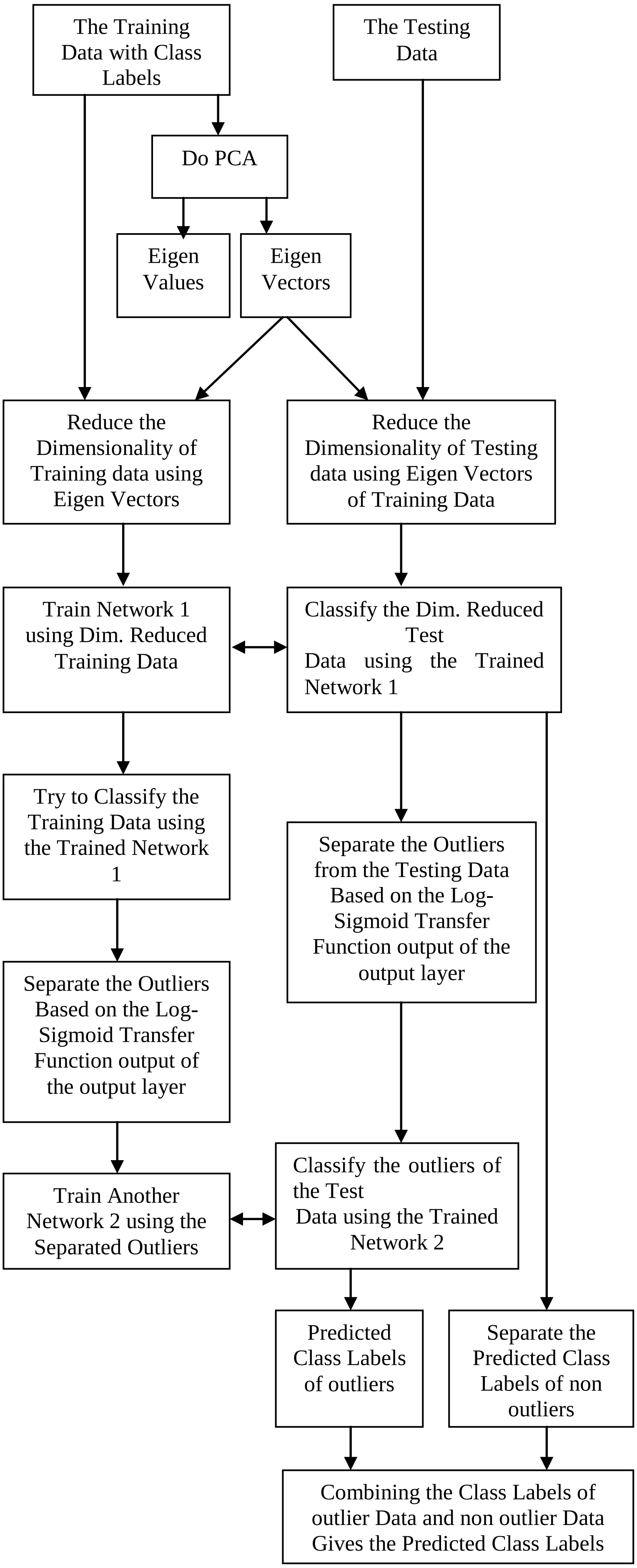


Figure 2: Proposed Improved Bi-Model NN Classification System.

The Above block diagram shows the proposed IBM-NN system. This model is almost similar to the previous BM-NN Model. But here use the Eigen Vectors of the training data to reduce the dimension of the training data as well as testing data.

The Steps in Improved Bi-Model classification System
Procedure IBM-NN {

1 Read training data D_L and targets C_L and test data D_T

2 $(\lambda i , u_i) \leftarrow PCA(D_L)$

Where λi – Eigen Values and u_i - Eigen Vectors and i : 1 .. n

3 $u_i D_L \rightarrow D_{LR}$ - the dimensionality reduced representation of D_L

4 Create Network N1 to learn D_{LR} and map it to the original output class C_L

5 Classify D_{LR} using the trained network N1.

6 Separate the Outliers O_L from D_L Based on the Log-Sigmoid output of the output layer of N1

7 Train Another Network N2 only using O_L

8 $U_i D_T \rightarrow D_{TR}$ - the dimensionality reduced representation of D_T

9 Classify the D_{TR} using the trained network N1.

10 Separate the Outliers O_T from D_T Based on the Log-Sigmoid output of N1

11 Find Class labels C_O of the outliers O_T using the Trained Network N2

12 Separate the Predicted Class Labels C_N of non outliers of D_T from step-9 output

13 Combining C_O from step-11 and C_N from step-12 Gives the Predicted Class Labels C_T of D_T

}

4. Result and Discussion

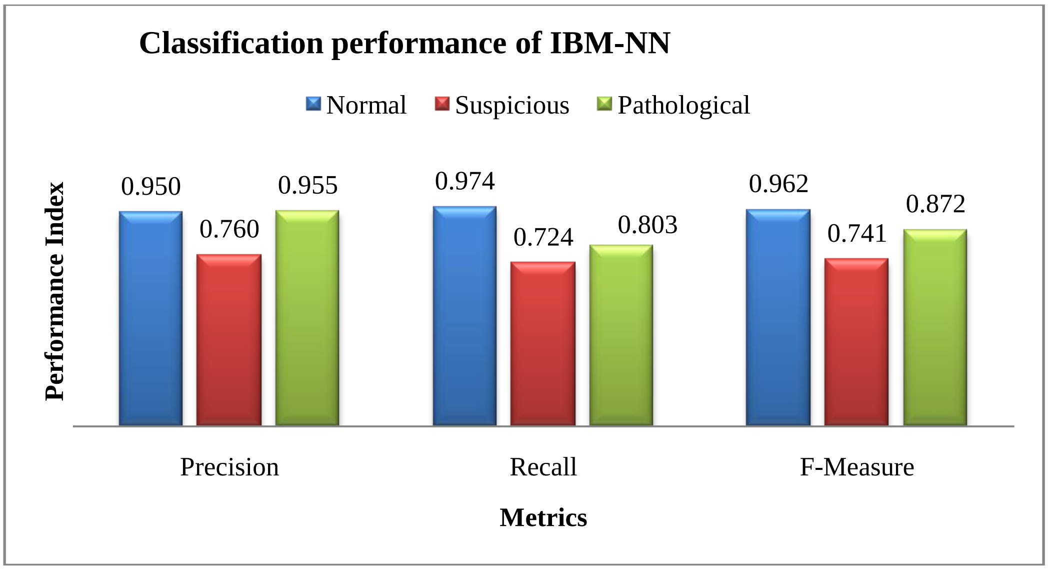
The following table shows the Classification performance of the proposed Improved Bi-Model neural network based classification algorithm.

Table 1 Classification Performance of IBM-NN algorithm

Class	Precision	Recall	F-Measure
Normal	0.9499	0.9735	0.9615
Suspicious	0.7603	0.7242	0.7411
Pathological	0.9551	0.8030	0.8718

The Analysis of Results

The following chart shows the Comparison of Precision under four different methods. The proposed IBM-NN based classifier provided good Precision in all the cases (Normal, Suspicious and pathological). Even though the performance of BPN in terms of Precision is good while classifying the Normal and Pathologic records, it was not good in identifying the suspicious cases. Particularly, the proposed method significantly improved the performance in the case of suspicious class.



The following chart shows the Comparison of Recall under four different methods. The Proposed IBM-NN based classifier provided good Recall in all the cases. In terms of recall, BPN was not good in identifying the suspicious and pathological cases.

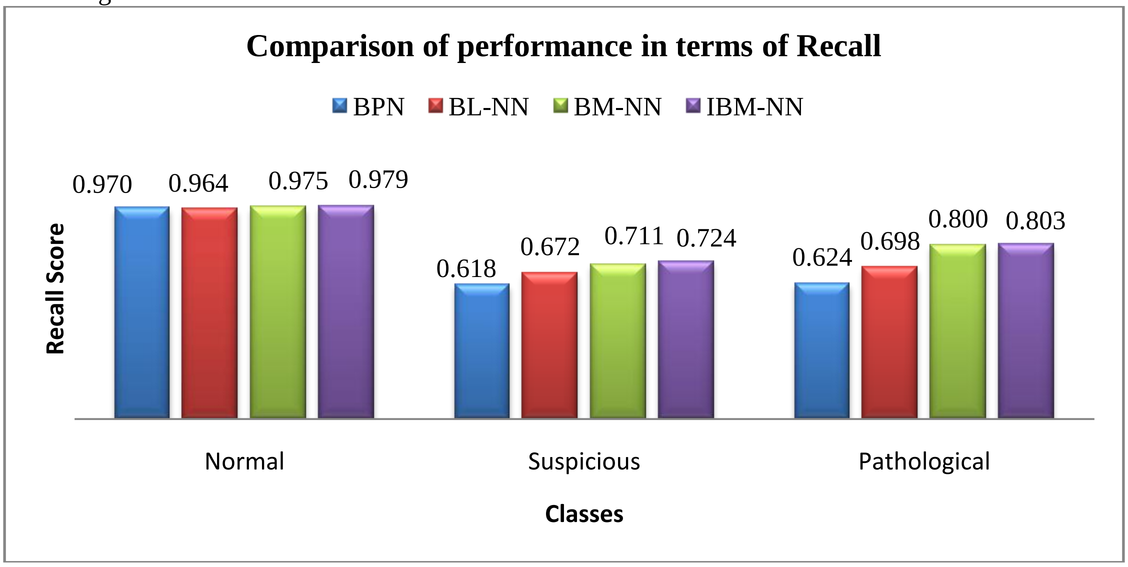


Figure 3: Comparison of Performance in terms of Recall

The following chart (Figures) shows the Comparison of F-Measure under four different methods. The proposed IBM-NN based classifier provided good recall in all the cases (Normal, Suspicious and pathological). Even though the performance of BPN in terms of recall is good while classifying the Normal and Pathologic records, it was not good in identifying the suspicious cases.

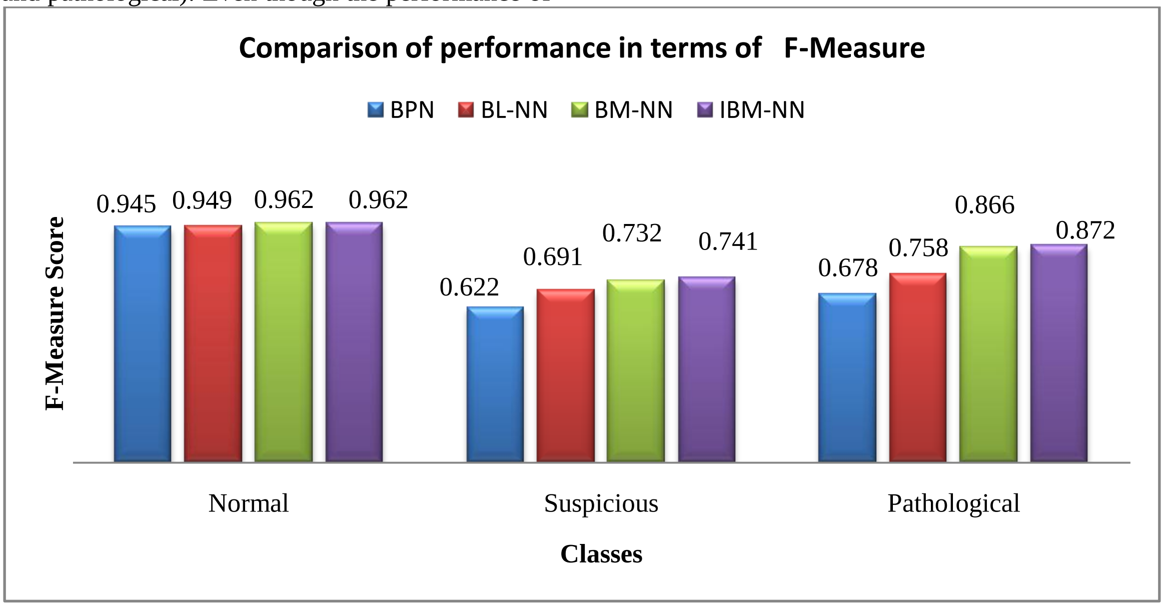


Figure 4 Comparison of Performance in terms of F-Measure

The following charts (Fig) show the performance of IBM-NN algorithm. In general, the algorithm gives good precision for normal records and poor performance in all other cases.

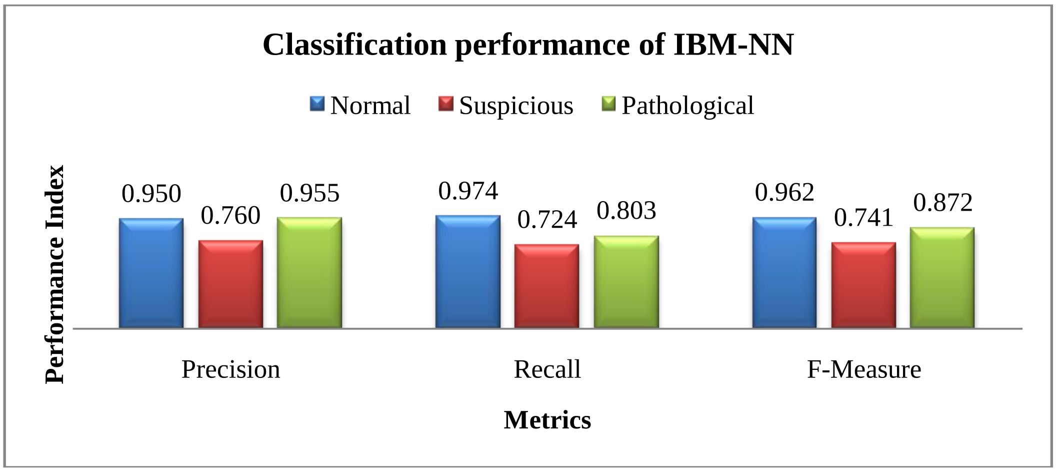


Figure 5 Classification performance of IBM-NN algorithm

5. Conclusion

We have successfully designed three hybrid neural network based classification systems and evaluated the performance of the methods with respect to four different metrics. The proposed methods IBM-NN were compared with normal BPN based method. According to the arrived results, the performance of the proposed supervised machine learning models provided excellent improvement in classification accuracy.

6. References

1. Sundar, C., M. Chitradevi, G. Geetharamani, 2012. Classification of Cardiotocogram Data using Neural Network based Machine Learning Technique, *Int J Comp App*, 47(14): 19-25. .

2. .Sundar, C, M. Chitradevi, G. Geetharamani, 2012. An Analysis on the Performance of k-means Clustering algorithm for Cardiotocogram Data clustering, *Int J Comp Sci & App*, 2(5): 11-20.

3. Sundar, C., M. Chitradevi, G. Geetharamani, 2014. Incapable of identifying suspicious records in CTG Data using ANN based machine learning techniques, *J Sci and Ind Res*, 73: 510-516.

4. Shahad Nidhal, M.A., Mohd. Ali1 & Hind Najah, 2010. A novel cardiotocography fetal heart rate baseline estimation algorithm, *Sci Res and Essays*, 5(24): 4002-4010

5. S.Anto, Dr. S.Chandramathi, “Supervised Machine Learning Approaches for Medical Data Set Classification – A Review” IJCST Nol. No.2, Issue 4, pp. 234 – 240, Oct – Dec 2011, ISSN : 2229-4333.